

```

1  AAGCGATAGC TGAGTGC GGCCTGCTGAT TGTGTTCTAG GGGACGGAGT
51 AGGGGAAGAC GTTTGCTCTC CCGGAACAGC CTATCTCAT CTCTTCTTTC
101 GATTACCCGT GCGCGGGAGA GTCAGGGCGG CGGCTGCGGC AGCAAGGGCG
151 GCGGTGGCGG CGGCGGCAGC TGCAGTGACA TGTCCAGCAT GAATCCCGAA
201 TATGATTATT TATTCAGTT ACTTCTGATT GGCGACTCAG GGGTTGGAAA
251 GTCCTGCCTT CTTCTTAGGT TTGCAGATGA TACATATACA GAAAGCTACA
301 TCAGCACAAT TGGTGTGGAT TTCAAATAA GAACTATAGA GTTAGACGGG
351 AAAACAATCA AGCTTCAAAT AGAGTCCTTC AATAATGTTA AACAGTGGCT
401 GCAGGAAATA GATCGTTATG CCAGTGAAAA TGTCAACAAA TTGTTGGTAG
451 GGAACAAATG TGATCTGACC ACAAAGAAAG TAGTAGACTA CACAACAGCG
501 AAGGAATTTG CTGATTCCCT TGGAATTCGG TTTTGGGAAA CCAGTGCTAA
551 GAATGCAACG AATGTAGAAC AGTCTTTCAT GACGATGGCA GCTGAGATTA
601 AAAAGCGAAT GGGTCCCGGA GCAACAGCTG GTGGTGCTGA GAAGTCCAAT
651 GTTAAATTC AGAGCACTCC AGTCAAGCAG TCAGGTGGAG GTTGCTGCTA
701 AAATTGCGCT CCATCCTTTT CTCACAGCAA TGAATTTGCA ATCTGAACCC
751 AAGTGAAAAA ACAAATTCG CTGAATTGTA CTGTATGTAG CTGCACTACA
801 ACAGATTCCT ACCGCTCTCA CAAAGGTCAG AGATTGTAAA TGGTCAATAC
851 TGACTTTTTT TTTATCCCT TGACTCAAGA CAGCTAACTT CATTTTCAGA
901 ACTGTTTTAA ACCTTTGTGT GCTGGTTTAT AAAATAATGT GTGTAATCCT
951 TGTGCTTTTC CTGATACCAG ACTGTTTCCC GTGGTTGGTT AGAATATATT
1001 TTGTTTTGAT GTTTATATG GCATGTTTAG ATGTCAGGT TAGTCTCTG
1051 AAGATGAAGT TCAGCCATTT TGTATCAAAC AGCACAAGCA GTGCTCTGCA
1101 CTTTCATGC ATAAAGTTTA GTGAGATGTT ATATGTAAGA TCTGATTGCT
1151 TAGTTCTTCC TTGTAGAGTT ATAAATGGAA AGATTACACT ATCTGATTAA
1201 TAGTTCTTCT ATACTCTGCA TATAATTTGT GGCTGCAGAA TATTGTAATT
1251 TGTGTCACAC TATGTAACAA AACAACGAA GATATGTTTA ATAAATATTG
1301 TACTTATTGG AAGTAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
1351 AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA
1401 AAAAA (SEQ ID NO:1)

```

FEATURES:

```

5'UTR:      1-179
Start Codon: 180
Stop Codon:  699
3'UTR:      702

```

Homologous proteins:

Top 10 BLAST Hits

	Score	E
CRA 108000024647144 /altid=gi 12728868 /def=ref XP_002675.2 RA...	372	e-102
CRA 18000004923424 /altid=gi 4758988 /def=ref NP_004152.1 RAB1...	332	5e-90
CRA 18000004937406 /altid=gi 131787 /def=sp P05711 RB1A_RAT RAS...	328	1e-88
CRA 18000004952860 /altid=gi 131785 /def=sp P22125 RAB1_DISOM R...	320	3e-86
CRA 18000004995539 /altid=gi 103720 /def=pir D38625 GTP-bindin...	313	3e-84
CRA 18000004967528 /altid=gi 92339 /def=pir S06147 GTP-binding...	297	2e-79
CRA 18000004880958 /altid=gi 464524 /def=sp Q05974 RAB1_LYMST R...	282	9e-75
CRA 18000004908714 /altid=gi 466171 /def=sp P33723 YPT1_NEUCR G...	253	3e-66
CRA 18000005175724 /altid=gi 7497231 /def=pir T33781 hypotheti...	253	4e-66
CRA 335001098696672 /altid=gi 11558649 /def=emb CAC17833.1 (AJ...	251	2e-65

BLAST dbEST hits:

	Score	E
gi 12867866 /dataset=dbest /taxon=960...	654	0.0
gi 12097820 /dataset=dbest /taxon=96...	654	0.0
gi 12793758 /dataset=dbest /taxon=960...	624	e-177
gi 12338056 /dataset=dbest /taxon=96...	622	e-176
gi 11977068 /dataset=dbest /taxon=96...	609	e-172
gi 10339840 /dataset=dbest /taxon=960...	517	e-145
gi 10349761 /dataset=dbest /taxon=960...	436	e-120
gi 10997958 /dataset=dbest /taxon=96...	385	e-105
gi 10996533 /dataset=dbest /taxon=96...	381	e-103

EXPRESSION INFORMATION FOR MODULATORY USE:

library source:

From BLAST dbEST hits:

gi 12867866	Fetal brain
gi 12097820	Adrenal gland
gi 12793758	Brain neuroblastoma cell line
gi 12338056	Adrenal gland
gi 11977068	Skin melanotic melanoma
gi 10339840	Uterus leiomyosarcoma
gi 10349761	Skin melanotic melanoma
gi 10997958	Placenta
gi 10996533	Placenta

From tissue screening panels:

Whole brain

```

1  MSSMNPEYDY LFKLLIGDS GVGKSCLLLR FADDTYTESY ISTIGVDFKI
51 RTIELDGKTI KLQIESFNNV KQWLQEIDRY ASENVNKLLV GNKCDLTTKK
101 VVDYTTAKEF ADSLGIPFLE TSAKNATNVE QSEMTMAAEI KCRMGPATA
151 GGAEKSNVKI QSTPVKQSGG GCC (SEQ ID NO:2)

```

FEATURES:

Functional domains and key regions:

[1] PDOC00001 PS00001 ASN_GLYCOSYLATION
N-glycosylation site

125-128 NATN

[2] PDOC00005 PS00005 PKC_PHOSPHO_SITE
Protein kinase C phosphorylation site

Number of matches: 5

1	59-61	TIK
2	97-99	TTK
3	98-100	TKK
4	106-108	TAK
5	122-124	SAK

[3] PDOC00006 PS00006 CK2_PHOSPHO_SITE
Casein kinase II phosphorylation site

Number of matches: 3

1	35-38	TYTE
2	106-109	TAKE
3	127-130	TNVE

[4] PDOC00007 PS00007 TYR_PHOSPHO_SITE
Tyrosine kinase phosphorylation site

30-36 RFADDTY

[5] PDOC00008 PS00008 MYRISTYL
N-myristoylation site

Number of matches: 3

1	21-26	GVGKSC
2	147-152	GATAGG
3	152-157	GAEKSN

[6] PDOC00017 PS00017 ATP_GTP_A
ATP/GTP-binding site motif A (P-loop)

18-25 GDSGVGKS

[7] PDOC00579 PS00675 SIGMA54_INTERACT_1
Sigma-54 interaction domain ATP-binding region A signature

14-27 LLLIGDSGVGKSL

BLAST Alignment to Top Hit:

>CRA|108000024647144 /altid=gi|12728868 /def=ref|XP_002675.2| RAB1,
member RAS oncogene family [Homo sapiens] /org=Homo
sapiens /taxon=9606 /dataset=nraa /length=222
Length = 222

Score = 372 bits (944), Expect = e-102
Identities = 190/222 (85%), Positives = 190/222 (85%), Gaps = 32/222 (14%)
Frame = +3

Query: 129 GGCGSKGGGGGGSCSDMSSMNPEYDYLFKLLLLIGDSGVGKSCLLLRFADDTYTESYIST 308
GGCGSKGGGGGGGGSCSDMSSMNPEYDYLFKLLLLIGDSGVGKSCLLLRFADDTYTESYIST
Sbjct: 1 GGCGSKGGGGGGGGSCSDMSSMNPEYDYLFKLLLLIGDSGVGKSCLLLRFADDTYTESYIST 60

Query: 309 IGVDFKIRTIELDGKTIKLQI-----ESFNNVK 392
IGVDFKIRTIELDGKTIKLQI ESFNNVK
Sbjct: 61 IGVDFKIRTIELDGKTIKLQIWDTAGQERFRTITSSYYRGAGHIIIVVDVTDQESFNNVK 120

Query: 393 QWLQEIDRYASENVNKLVLGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ 572
QWLQEIDRYASENVNKLVLGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ
Sbjct: 121 QWLQEIDRYASENVNKLVLGNKCDLTTKKVVDYTTAKEFADSLGIPFLETSAKNATNVEQ 180

Query: 573 SFMTMAAEIKKRMGPGATAGGAEKSNVKIQSTPVKQSGGGCC 698
SFMTMAAEIKKRMGPGATAGGAEKSNVKIQSTPVKQSGGGCC
Sbjct: 181 SFMTMAAEIKKRMGPGATAGGAEKSNVKIQSTPVKQSGGGCC 222 (SEQ ID NO:4)

Hammer search results (Pfam):

Model	Description	Score	E-value	N
PF00071	Ras family	256.4	7.7e-75	2
CE00060	CE00060 rab_ras_like	170.0	3.9e-47	2
PF00634	BRCA2 repeat.	9.9	0.39	1
PF00056	lactate/malate dehydrogenase	3.9	3.4	1

Parsed for domains:

Model	Domain	seq-f	seq-t	hmm-f	hmm-t	score	E-value
PF00056	1/1	13	29 ..	1	18 [.	3.9	3.4
CE00060	1/2	8	64 ..	20	77 ..	86.8	8.9e-23
PF00071	1/2	13	64 ..	1	52 [.	111.9	4.8e-32
PF00634	1/1	57	79 ..	13	35 .]	9.9	0.39
CE00060	2/2	65	140 ..	110	188 ..	81.2	2.9e-21
PF00071	2/2	65	173 .]	85	198 .]	142.4	4.5e-41

1 TTTTGGGTGT GTGTGTGTGT GTGTGTGTGT GTGCCTTTAC TAGTGA CTCA
51 GGTACACAGTT TTCTGAGATT TTTTCTCTCC CCTCAAGACA GAATCTTGCT
101 CTGTCGCCCA GGCTGGAGTG CAGTGGCCTC TCGGCCCACT GTAGCCTCCG
151 CCTCCCGGGT TCAAGCAATT TTCCTGCCTC AGCCTCCCGA GTAGCTGGGA
201 TTACAGGCAC GCGCCACCAT GCCTGGCTAA TTTTGTATT TTTAGTAGAG
251 ACAGTGT TTC ACCATGTTGG CCAGGCTGGT CTTGAATTCC TGACCTCGTG
301 ATCTGTCCGT TTTGGCCTCT CAAATTCCTG AGATTACAGG CATGAGCCAC
351 CGAGCCTGGC CAGTTTTCTG AGTTTTTATT TGAATCAAA ATAAGCTTTT
401 TTTTTTTTTT TAATGGGCTT TAGAGTCCAG GGTAACGAAC ACTTTTTGGT
451 GCCTATTACT GAACCATTCA GGGTATTCCT GGGGTGGTGA CCGTGTTTCA
501 TTCAGAAACC AACATGTTCA TTTCAGAAAC CAACTCGGG TAACTTTTGA
551 TAAGTTCATC AACTAAGGCC CATGGCAGAA TTTGAGGGCT AAGGGGTGTA
601 ATTAGTGTAT GGGTAGAAAT AAGTGCCTTC TTTCTATATT TTGGCGTTGT
651 AGGAATTTAA AGTGATTCTG CAGTAAGTCT CAGGAGACAA TTTTCTTAGT
701 TCTTAGAAGT TGGAAGATAA ACTTTGGACA ATGTATTACA CTATGCCCTT
751 TGTAAATAAA TAACTCAAGA TAATGTGTTA AAGTTTAGCG GAGATTTAAA
801 TTCCTGAGCT GATTAAAGAG AGCTGTTAAG GCCATAGGTT TTTTAAAAAT
851 GAGTTAATAT TACTCCCAGA AATTGTAGGC ACTATATAGT GATGAATTGC
901 ATATTTTTTAT TGCTTATTAT TTTCCAGTCT TGCAGAATGG CTCAGGGTTA
951 GTAGCAACTA AAAGATAATA CATTACAATT CAACCTGAAG GCCGGGACGA
1001 AGGTAGGAAT TGGATTTTAG GCTGGCTCTG GGCTGTGTCC CTCCCATCCA
1051 TGGGATGTGG AGCCATTGAA GGTGTGGGG TCACGATGCA GGTGCTGTCT
1101 CAGAAAGATA CATCCGACTG TGTGTGCAAA TGGGCTGGGG CGGAGAAGAG
1151 AGAGAGAGGT AGAGTCCATT TGGAGACTAC TGCAATAGCC AGGCTGACGA
1201 GTTAAGAGCG GGGCACAGTA AGAATGGGAA GAAATCTAAG AAGAAAATGG
1251 TAGTGCGCGG GGGCAACAAT GGACGATGAC CGAACCAGG TGGGGATGGG
1301 TGAGTGACGA GAAGAACCGC TCCGTGCCGT CCAGGGAGCC CTTGACTTC
1351 CCTTCTGTTT TTAGAGCGGA CGTCTCCTA CCAGCCCCCA ACCAGCGCCA
1401 CCAGGGTGGC GCAAGCCTCA AGCTGGTCAG GTCAGCAACA GCCGCAACGG
1451 AGGCAGGAGC CGACACGCTC GTACCCCGGC CCCCTCCCG CCCCGCACCC
1501 CCCGGCAGTC CCTCCGTTT GACCACTCCC CCCGGTCCCT TGCCTCCCCC
1551 GACCCCCAGC CTCCGTGGGC CGCCGGCACC ACCCTCCGCC CCTCTCCGCC
1601 CCTCCCCCG TGGGGCGCTG ACTCGCCCGG CTGCCACGTC TCACTGATGA
1651 CATCACTAGG GCAGCTCGGC CTTAGCCAAT CCGCCAGGGG GAGTCCGAGC
1701 GAAGTCTAG CCAGCGAGTC AGAGGGGAGG GGAGCAGGGA GGGGCCGAGG
1751 GTGGGGAGGT GAGGGAGTGG GGAATGGGGC GGGCGACAAC CCTTCAGGTA
1801 CGCATGCCCC AGAGGCGCGG CGCTTGGCGG GAAGCTGAGT CCTGGCCTTG
1851 CGTCGCACTG TCTGTCTCA GCTCGCGTAG CCGCGCTCGC GACTCCCTTT
1901 CCCGGCATGC CAGGCGGTGC GGCCGCCCTC TGGGCCGTGT AAAGGCCCTT
1951 CGGTCTAAGG CTTCCCTATT TCCTGGTTTC CCGGCGGCCA TTTTGGGTGG
2001 AAGCGATAGC TGAGTGGCGG CGGCTGCTGA TTGTGTTCTA GGGGACGGAG
2051 TAGGGGAAGA CGTTTGCTCT CCCGGAACAG CCTATCTCAT TCCTTTCTTT
2101 CGATTACCCG TGGCGCGGAG AGTCAGGGCG GCGGCTGCGG CAGCAAGGGC
2151 GCGGGTGGCG GCGGCGGCAG CTGCAGTGAC ATGTCCAGCA TGAATCCCGA
2201 ATAGTGAGTT CAGGAGAGCA CCGGTCGGCT GGGTCCGTGG GCCAGCTTGG
2251 GGGATCTTAA AGGGGTGAG GAGGGTTGGG GCAGAAGTCG GGGCATCGGC
2301 TGGGGTGAGG CGAGGGTGAT GGTTCAGGAG AGGCTGGCGG CCGGGAGTCG
2351 GGCCCATTTG TCTGACGCG AGGGGCGGCC GCGCGGGGGA GGGGTGCGGC
2401 CGGAGGGGTG AGCCGCCCG GCCTGGACCG GGTCAAGTTA GAGGGCTGA
2451 CTGCGGGGCG GGTGCTGAGG AAGCCTGCCG AGGGGCCCTG GCGGTGTGA
2501 AGGGGTATCT TCTCTCGGAG GCAGTGACTT TTGAAGGAGG ACTTGTCTCT
2551 AAGGGGAGGG GATGGGGTGG GAGAGCCCTT CTAGAGGGCA CTGTCAGACC
2601 CTGCGCCCGC ACTCTGCGGA GCTGTCAGGA TCTTCGGGGT AGAAACCAGC
2651 TTTACTTGTA AATCCTGAGC TTGTTGGGTC TCTCTCCTTC CATCCTCCCC
2701 GCCAGGTTT CAGGTAATAG GATGCTTTT GGGACTGCGT GGGATTGAGG
2751 GGAATGAGTA GATGGTGAGA AGCAACTGAA CATTATATTAG TTCTCTTTT
2801 GAGTTGTGTC TTGGAGGAGT TGTTTAAGAG CTCGCCGGGT CCATTGCCCT
2851 CCTATAAAAA CCTGGGCATT TGTGAGAAAT TTGTTTTTTT TTTTTTTAAA
2901 GAGGACACCT AAGTCATTTT GTCTTCTGTG GGTCAAGGGA AAAAAAAAAA
2951 ACTAAAGCCA AGAAATGTCT TTTTGATACT CGCAGATTAA AGGAAGCTTG
3001 CTGTCAAGTT GAAAGAGAAA CGAACGGGAC CTATGATAGA TCTGTATGTA
3051 GGTTTTGGAT TACCTGCTTG GATGCTTGCA GATAGGGAAT GAGGTTCCAT
3101 GACGTGTCAT GAAAAGTTAA TGCATTTCTT TTTCTTGCTT ACTCAAGAAG

FIGURE 3, page 1 of 21

3151 TCACCACAGC AGATGTGACA CACCTGGCAC CTTTCCTGGG AACTGGTGT
 3201 CACTTCCCTT GGGTAGAGTT TGGTGGGCTC TCCTCAATGG CCCTTTAAAA
 3251 ATTTCTCTTA CAGTTTACAT GCATGTAAAG TAATGAATAA TTGGAAGAGA
 3301 CCGAATTGGT ATTCCTTTTC AGTGTCAAAG GCCTTTGAGG GATGGGGGAA
 3351 AATCAGTATT TGTGTAAAA GTTGAGTTTA TTTGCTGGTT TGGTCAATTA
 3401 CTGCTAGACA TTTTCCCTTA AAAGGTCCAC CCACCAGTTT AGCTGACTGT
 3451 CATATGTGTG TCACATGGCT CTTGCAAAAT GCTTACAAGT TTTGTAATAG
 3501 TGTGGCTTGA AGCTGAAATC TTTTGCACTA AACAGAAACC GTAGTATTTT
 3551 ATTAGAATTT CATGCTTTAG AAGTTGAGGG TAGTGTCTT GTAGTGACAT
 3601 TTGCTGTGTT GACAGTTTAA AAAAATTTTT TTTTCAAGGG CTCCAAGGAC
 3651 AAAGTTGGTT TTGCACAGTT GAACGGAGGT GAACCTGAGG TTCTTAATTT
 3701 AGTAGTTTTT TTGGTAACAA TAAAGAACAT GGATTACTG CTTTATCGAG
 3751 GTTTATAGAC CTCTACTGTT CAGGAAATTT TCTGAATTTG CTATATATAT
 3801 GTTTATTAGT GTAAATAAAT CTTCAAGATT AGTTGAGAAC TTTGACAAGT
 3851 TACTCAGCCT CTGAATTTTT TTTCCCTTTT GTAAAATAGG ATAATTGGAG
 3901 TCATTATTCC TGTGAGGTA GTGGTGAAT TCAAATGTAT ATAAAAGAAT
 3951 TTGAAAAACT GTGTGAGCAT TCTTCAGGTG GTATGCATCA TTTTCATGAA
 4001 AGGCATTCTA TTAGTACCAG GATTTAGGAA TATAATCCTT GCGCTTAAGA
 4051 AGTTTAGATA TAGGCCAGGC GCGGTGGCTC ACCTCAGTAA TCCAGCACT
 4101 TTGGGAGGCC GAGGCGGGCG GATCCCGAGG TCAGGAGATC GAGACCATCC
 4151 TCGGTAACAC GGTGAAACCC CGTCTCTACT AAAAATGCAA AAAAATTAGC
 4201 CGGGCGTGGT TGTAGTCCCA GCTACTCGAG AGGCTGAGGC
 4251 AGGAGAATGG CGTGTATCCG GGAGGTGGAG CTGTCAGTGA ACCAAGATCT
 4301 GGCCACTGCA CTCCAGCCTG GACGACAGAG CAAGACTCCG TCTCAAAAAA
 4351 AAAATTATTT ATTGTTTTGA GACGGAGTTT CAATCTTGTT GCCCAGGCTG
 4401 GAGTGCAATG GCGCAAACTT CCTCTCACC GACCTCCGC CTCCTGGGTT
 4451 CAAGTGATTC TCCTGCCTCA GATTCCTGAG AAGTTGGGAT TACAGGCATG
 4501 TGCCACCACT CCCGGCTAAT TTTGTATTTT TGGTAGAGAC GGGGTTTCTC
 4551 CATGTTGGTC AGAATCTGCT CAACTCCCG AAGTGATCCG CCCGCCTCAG
 4601 CTTCCTCAAAG TGTGAGGATT ACAGGCGTGA GCCACCGCG CCGGCAGAAA
 4651 TAGATTTTAT ACATGTCAA TACCAGTAGA TATAGCAAAT TCCAGATGTG
 4701 TGGCATGGAT GAGAGCAACA AGATTTCAGG GGGATGGTGG GTTGTGGTTG
 4751 GCTATCTGGG TTTTGGAGA CTTTATAGAA GAGAGACCTG AAAGGGATT
 4801 ATCAGCAATT AGATTGGAG GAACAGAGG AGTGACTAGG AATTTTCAAG
 4851 GGGGAGAAGA AGGAGGAATG GCTCATAAAT GACAAGGACA GTAATAAGTA
 4901 AATACGGTGT AGAATCATCC TTTCTTTTGA AGACTAATGA CCTCAAAGGG
 4951 ATCAAACCCA GAAACAGTTT TTATATTTTT TCTGGGATCA AATACATGGG
 5001 TATCTGGCCT ACTATATTTG TATTCTAGAC TGTTTAGTAA AATAATACAG
 5051 GAATTTGAGA AAACCTTTGC AAAAGTGTTA GTGAAAATTA CTTAGGGTGA
 5101 GAGGAAGTGA GGGATATTTT ATTAGGGGAG GTCACAAGGG CAGTGAGCAA
 5151 TCAGATTTTT AGTAATCTGA CTTAAGCAGT TTCTTTTTGT TTTAATGAAG
 5201 CTTGTTTATC TTATAAAGT AATTAGAGAA AATTTGGAAA ATAAAGGAAA
 5251 GAAAGAAAAA TTTCTTAGTG TTTTATCACG CAAATACAAG CTCATTCTGT
 5301 TTTAACATCT TGTTCCAAAC TCCAAAGTCT TGCTTTCTCT TCAATTAAAA
 5351 CTTTAATGGG TGGATGCTTT TCCTGCTTCC AGTATGTTAT CTTAATAACT
 5401 AACAATGGTA TATTAGCTAA TGTTACAAA TGTACTCCAG ATGTTCTTAA
 5451 AGTTACTTTG GTTTATCATT ACCAATTTAT ATTGTTTCTT TTAGAAATTT
 5501 ATAATCTTTG TTAATGGGTT CTGCTAAATT TGGTAGTGAA AATGGGATCT
 5551 TGAGAAAAAA GATTCTGAAG CAACAGAATT TTTAGATTTA TATTGGTTTA
 5601 CATAAGAGTT GGTAGCTGTA TTACTTTTTT TGTTTGTTTT GTTTTTTTTT
 5651 TGAGACGGAA TCTTGCTCTG TCGCCAGGC CTTGGCCTCC CAAAGTGTG
 5701 GGATTACAGG CGTGAGCCAC TGTGCCTGGC TGTTTGTTGT TTTTTTTGTT
 5751 TTTGTTTTCT TTTCTTTTTT TTTTTTTTCA GATGGAGTCT CACTCTGTCA
 5801 CCCAGGCTGG AGTGCAGTGG CGCGATCTTG GCTCACTGCA ATCTCTGCCT
 5851 CCTGGGTCA AGCATTTCCT CTGCCTTGGT CTCCTGAGTA GCTGGGATTA
 5901 CAGGCATTGG CCACATAAC CAGCTAATTT TTGTATAGAG TACCCAGCCA
 5951 TCTCTAATGT TGATCAGGCT GAAGCAGGTG GATCACCTAA GGTCAGGAGT
 6001 TCAAGACCAG CCTGGCCAAT ATGGCAAAAC CCTATCTCTA CTAATACAGA
 6051 AAATTATCTG GGTGTGTTGG CTGGCGCCTG TAATCCCAGC TACTCGGGAG
 6101 GCTGAGGCAG GACAACTCTT TGAACCTCGG AGGTGGAGGT TGCAGTGAGC
 6151 CGAGATCACA CCATTGCACT CCAGCCTGGG CAACAGAGCA AGACTTGTCT
 6201 CAAAAAAGAA AAAAAAAGGC AATTGAAAGT GTAATCTGAA
 6251 CAGTTAAAAA AGTAGATAGA AAGGGTTAA GCTTTTTTTT GAGGATCTGA

FIGURE 3, page 2 of 21

Figure 3, page 3 of 21

6301 AGAAAAATGT GGATTTTTTTT TGAGCTACGT TTTGAAGCAG GCAGTGATTA
6351 TTTACGACACA TTAAGAAATG CTTAACATGG CCAGGCCGAG TGGCTCACGC
6401 CTGTAATTCT CAGCACTTTG GGAGGCCGAG GTGGGCGGAT CATTGTAGGT
6451 CATGACCAGC CTGGCCAACA TGATGAGACA CTGCCTCTAC TAAAAATACA
6501 AAAATTAGCT GGGTGTGGTG GTGCACGCCT GTAATTCAG CTA CTACTCAGGA
6551 ACCTGAGGCA GGAGAGTCAC TTGAACCTGG GAGGCGGAGG CTGCAGTGAG
6601 TCCAGATCAT GCCACTGCAC TCCAGCCTGA GGGACAGAGT GAGACTCCTC
6651 AAAAAAAAAA AAAAAAAAAA AAAGAAATAC TTAACATTAT TCTCGTGATT
6701 ATTCTCATAA CATTTTTTCAT AATCCACTGG CTTCCAGTGG ATTTTTTTAG
6751 TGTCAAGAAA ATAATTTTGA TTGGTTCATC TTTAAGGAAT GTGTTAAGAA
6801 TAAAGCATGT CTACCTGTCT TCAGTATACC AGCTAACTAT AGTAGGAAGA
6851 AATATAGTAG TCTACTTAGA TCAACTATAA TTCTTTAATG CAGAAAAAGT
6901 TTAAAGTATT TACCTTATT TTAGCCCCCA TCCCCTTAAG TATATCATGG
6951 CTCCAGAATC TCTGAAAATG TTATCAGTCT TTCAGACTTT GCTCTTCTTT
7001 CATGTTATAC TCAAGAAACA TTTGACCTTT TTTTTTTTTT TTTTGCTTGC
7051 ATGTGTGTTT AAATAATTTT TAACAAAAC TAAGTGTGTA AAAGTGAAAG
7101 CAGGTTGTCT TTGTGACTTT TGGTGGTGGT TTGAAAAACT CAGAAAAAGT
7151 TAAAGAAGAA AGATAACTAG TATTTCTCATT GTCCAGAATA TGATTTTTTA
7201 AATGTCTATA GAATATCACC ATCTGTAATT CTTCCGGTAA TTTAAGTATT
7251 CAGTAGTTGT ATAAAACCTT TAAAAATAT ATATTGAGAA TTTTGTGTGA
7301 ATGAGATGAT GAGATAATCT TG TAGGATCA TTTAAAGATA AGAACTGAGG
7351 CCTGGCACAG TGGCTCATGC CTATAATCAC AGCACTTTGG GAGGCCCAGG
7401 CGGTAGATCA CCTGAGGTCA GGAGTTTGAG ACCAGCCTGG CCAACATGGC
7451 AAAACCCTGT CTCTACTAAG CATAGAAAAA TTAATTGGGT GTGGTCGTGC
7501 CTGCGTGTAG TCCCAGCTGC TTGGGAAGCT GAGGCGGGAG AATCTCTTGA
7551 ACCCTGGAGG TGGGCATTGC AGTGAGCTGA GATTGCGCCA CTGCACTCCA
7601 GCCTGGGCGA CAGAGCAAGA CTCTGTCTCA AAATAAAGTA AAATAAAATG
7651 AAGATAACAA CTGAAATTTT ACATTAATAA TTTTTTTGTA GCGACTGTGC
7701 CTCTATGTT GTG CAGGCTG GTCTCAAAC CCTGGCCTCA AGCGATCCTT
7751 CCAAAGCACT GGGTGGGCCA CCATGTCCAG CCTGAAATTT TGCATTAAAA
7801 AATTTCCCGC TTTTGGCTGG GCGAGGTGTC TCACGCCTGT AATAGCAGTT
7851 TGGGAGGCCG AGGCAGGCAG ATCACTTGAG GTCAGTTCTA GACCGGCCTG
7901 GCCAATGTGG TGAAACCTG CCTCTACTAA AAACACCAA TTAGCTAGGC
7951 GTGGTGGTGT GCGCTTG TAG TCCCAAGCTA CTGAGGAGGC TGAGACAAGA
8001 GAATCGCTTG AATCTGGGAA AAAGAGGTTG CCGTGAGCCA AGATTGGCCA
8051 CTGCACTCCA GCCTGGGTGA CAGAGTGAGA TTCTGCTCA AAAAAATAA
8101 AAATAAAAT TTCCCCCTT AATCAAATTA AGTTAAATG AGGGATGTTA
8151 GACAGTTTTT AACCATCAA TATTTTAGTT TAGTTTTTTT TTTTAACTG
8201 TGTCTTAAAG ATGGAAGTGC TTCAAATCA AATCTTCCTT GCCAGTTCTC
8251 TACTTGGCTT CTTTTTTTTT CTTTTTGAGA TAGAGTCTCA CTTTGTCACT
8301 GGAGTGCCTT GGCCTGATCT CCGCTCACTG CAACCTCCGC CTTCCAGGTT
8351 TAAGTGATT TTCCACTCA GCCTCTCAAG TAGCTGGGAG TACAGGTGTG
8401 TGCCACCACA CCGCGCTAAT TTTGTAGTT TTAGTAGAGA CAGGGTTTCA
8451 CTATGTTGGC CAGGCTGGCC TCAAACCTCT GACCTCGTGA TCCACCCACC
8501 TCAGCCAAAT TGCTGGGATT ACTTGTGTGA GCCACGCGCC TGGCTTCTAC
8551 TTGGCTTTTA AAGGGAATTT TGCTTTCTGA GTAATTTTAT TTCTCAGGTA
8601 TCTTGGTCTT TTTAATTCTG GAAGCAATCT TAATAATTTA TGTATGTGCC
8651 CTGTAATCCC AGCACTTTGG GAGGCCGAGG TGGGCGAATC ACGAGGTCAG
8701 GAGATCGAGA CCATCTGGC TAACACGGTG AAACCCCATC TACTAAAAAT
8751 ACAAAAAATT AGCTGGGCGT GGTGGCAGGC GCCTGTAGTC CCAGCTACTT
8801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
8951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9001 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9051 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9101 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9151 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9201 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9251 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9301 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9351 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
9401 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN

FIGURE 3, page 3 of 21

FIGURE 3, page 4 of 21

12601 GCTGGGATTA CAAGCACCCA CCACCATGCC CGGCTAATTT TGTGTATTTT
12651 TAGTAGAGAC TGGGTTTCAC CATGTTGACC ACGCTGGTCT CGAACTCCTG
12701 ACCTCAGGTG ATCTGCCTGC CTTGGCCTCC CAAAGTGCTG GGATTACAGG
12751 TGTGAGCCAT CACACCAGGC CTCAAGAACT TTTTATTTTT GAGACAGGGT
12801 CTCACTCTGT CACCCAGGCT GGAGTACAGT GGTGAGATCA TGGCTTACTG
12851 CAGCTGGAC TCCCAGGCT CTGGTGATCC TCCCATCTCA GCCCCTGGAG
12901 TAATTAGGAA TATAGACACA CACCCATGCC TGGCAGTTTT TGTATTTTTT
12951 TTCTTTTTTC TCTTTTTTTG TAGAGACTGG GTTTCACATG TTGTATCAGG
13001 CTGGTTTTGA ACTCCTGAGC TCAAGCAATC CTCACTCTTT GACCTCCCAA
13051 CGTGCTGGGA TTACAGGCAT GAGCCACTGT ACCTGGCCTT TTCTACATTA
13101 AAAACTTTTT ATTAATAAAC CCAAATCTTC CTTGTGGTTG TATATACATA
13151 TATACATAGG TACACACATG GAGAATTTTA CCTTGGAGGA AGGCTTGGTA
13201 AAGAAAATAG CCCTTTGGGC CGGGTGCGGG GGCTGACGCC TGTAGTCCTA
13251 GCACTTTGGG AGGCTGAGGT GGGCGGATTG CCTGAGCTCA GGAGTTCAAG
13301 ACCAGCCTGG GCAACACAGT GAAACCTGT CTCTACTAAA ATACAAAAAA
13351 TCAGCTGGGT GTGGCAGCAT GTGCCTGTAG TCCCAGCTAC TTGGGAGCCT
13401 GAGGCAGGAG AACTGCTTGA ACCCGGAGG CAGAGGTTGC AGTGAGCCGA
13451 GATTGTGCTA CTGCACTTCA GCCTGCGCGA CAGAGCAAAA CTCTGTCTCA
13501 AAAAAACAAA CAAACAAACA AAAAAGGAAA ATAGCCTTTC TCTATCATCA
13551 GAGTATATTA AGAGTTGAGT TTTTTTTTCT GTTTTTTAA ATTTTTGTTG
13601 TTTATTTTAA ATTACAAAAC ATGGACTCTG CTTACAAATT AAGAAAATGA
13651 CTCATGTTCA AACAAGCATA ATCAATATAA CAGTTAATAC AAGTTAAATA
13701 TTGTAATATG TTTACGGAAT AGCATGGCAA AATAGTGCAA AAGATTGGG
13751 GAAGGGGCCT ATAATTTCTG TTAACAGAAA GTTTTAGTTA TGTGATTCA
13801 ACTGGAGAGG AACAGAGCTC CCAGAAGGAC TCCAGAACAC TTGATGCTTG
13851 TCTGAGTGGG GTCAGCAGCA CTGAGTTCCC ACCAGCCAGA AAGTTTGTGT
13901 GTGTACATTA TTTCCCTTAA CTGCCACAAT AATCCCATGA AGAAAATGCC
13951 CTAGTTTAC AAACAAGGAA ACAGAGGCAG AGAAGAGTTA AATGACTTGC
14001 CCAAGGCCAT TCAAAGTAAG CAACTGAATT GGAATTTTAA CTCAAAGGCT
14051 TGGATGTCCC ACTACAACAA ATAGGCTGTT TCTGCTTAC TACATGTGCT
14101 TACTTCTAAG AATTTAACAT TTTAGGCTGG TTGTGGTGGC TCACTCCTGT
14151 AATCTCAGCA CTTTCGAGG CTGAGGTGGG TAAATCACTT GAGCTCAGGA
14201 GTTTGAGACC AACCTGGGCA ACATGGTAAA ACCTCATCTC TACCAAAAAA
14251 AAAAAAAAAA CTAGCTGGAC GTGGTGGCAC GCGCCTGTGG TCCCAGCTAC
14301 TCAGGAGGCT GAAGTAGGAG GATCGTTTGA GCCTGGGAGG TGGAGTTGCT
14351 AGTGAGCCCA CATTGCAATCA CTGCACTCTA GCCTAGGTGA CAGAGTGAGA
14401 GCCTATCTCA CACACAAAAA AAAGAATTTA AAATTTTAGT CAAGTAATTA
14451 GGCATAACA TTTTGTGGTC AGTTACTTTA CGAATTCATG GTTGGAGGCC
14501 TGATGTGGTG GCTCATGCCT GTAATCCCAG CACTTTGGGA GGCTGAGGCA
14551 GGAGGATTGC TTAAGGCCAA GAGTCAAAAT CAGCCTGAGC AACCTAGTAA
14601 GATCCCCCTT CTGCAAAAAA TTTAAAAATT AGCTGGGCAT GGTAGTGTGC
14651 ACCTGTAGTC CCAACCACTT GGGAGGCTGA GGTGGGAGGA TTGCCTGAGG
14701 CCAGGAGTTT GAGACCTGGG CAGCATATGA AGACCCTGTC TCTAAAAAAC
14751 TAAAAATAAA AAATAGCCAG GTGTGGTTGG TGTGCTTGTG GTCCCAGCTA
14801 CTAAGAGGC TGAGGCAAGA GGGTTGCTTG AGCCCAGAAG TTGGAGGCTG
14851 CCGTGAAC TGATTGCACC ACTGCACTTC AGCCTGGGTG ACATAGCAAG
14901 ACCCTGTCTC TGTGGTGGTG GTGGGTGGGG GTGGGGGAAG GGATTTAAGA
14951 AGGGTTTGTG AGGTATGTAT TATTTATAAA TGGGCTTTTA ACTTTACCCT
15001 TCACATCTTG GGTGAAATT AATTGTATCC ATTCTCAGT TTTCTGTCTT
15051 GCTATATATT TAAACTTGGA GACTTAGAGG TCATGGATGT CTTTCTATGA
15101 AAAGCAAATG AAGCAGAGGG CTGCCTTCTC TTGCTGTAGA GGGCACACTT
15151 GCTGCAGAGC ATGTTACTGT TTTATGCATT GCTAGGCTTT GGGAGTTGTG
15201 ACTTGTATGA TCATAGTACT TACAACATAT AGTTGGCAAT TTTTAAACTT
15251 TAACTTTAGA TTATATATGT AAACCTCTGT GTTCCTTTGT CACTGATAAT
15301 CTGAACAGAA CCTTGGATA AATAATTTTG AAGTTTTTGT CTGAACCTCT
15351 GAAATTTGTA TTGTTATCTC ATGGTTTTGC TGGGAGGAAG GAGAAATAAC
15401 AATGGCCACT TACTGTGCTT CTGTATGTGC CAGACAGTAT GTGCTAGATG
15451 TTTCAGAAAC GTGATTTGTA ATCCTGACAA GAAGCCTAAT TGGGTGGTAG
15501 TGGGTGCTAA TTGAACCTTA TAGATGAGGA AATTGAGGCT CATGGTGGTA
15551 AGTGAATAAC TTGCACCAAG ATCCTATGGC TGGTATGCAG TAGAGCCTCA
15601 ATTCAAGTAC GGTCTCTCCA GGTCCAAACC CATGCAGGCT TTGAGAGGTA
15651 AGGAGGTAGA GAACGTTGAC ACCCCTTCT TGGTGTGTTT TTCAGCAAAT
15701 ACTTGTATGC ATATTAAAGA CTGTCTACCC TTTTGTCTATC TTGTGTCAC

FIGURE 3, page 5 of 21

15751 TGCTGCTTCC TTTGGTACTA CCCAAATTC TTTTCAGCATT TCAGCTTTGA
15801 ATTTTATTTT TTATTTTATT TAATTTATTT ATTTTTTTGA GATGGAGTCT
15851 CACTCTGTTG TCCAGGCTGG AGTGCAGTGG CGTGATATCA GCTCACTGCA
15901 ACCTCTGCCT CACAGGTTCA AGCAATTCTT CCTGCCTCAG CCTCCTTAGT
15951 AGCTGGGACT GGAGGTGCCC ACCACCACGC CCAACTAATT TTTGTATTTT
16001 TAGTAGAGAT AGGGTTTTAC CTGTGTGGCC AGGCTGGTTT TGAACCTTTG
16051 GCCTCAAGTG ATCCACCCAC CTCGGCCTCC CAAAATGCTG GGATTACAGG
16101 CATGAGCCAC TGCACCTGGC CAGCTTTGAA TTTTITAGAAT ACTGTTCTAA
16151 ACAGAACTAT ATTGGAACCT GGAAAATTAA TCTATTGTCT CTAAATACCA
16201 AAGAAAAACA TGTAATTTTA GTGGTTGATT ATGGGAACAA TTTTITTTAA
16251 GATGGTTCAT CTGAATGGGA AGCATTTTTTT TTTTAATTGC TTGACTATTT
16301 CTTTAAATTT GGAGAAAAGA CCATTGCCCT CTCAGATTTT TGGTAATTGG
16351 TCACATTGAT CATTTATATT GACTGACAGG CTGCTTTGTC CACAGCTGAA
16401 GGATTGTTTA ATTTTITTTA AATTATAAGA GTAATATGTG CTCACTGTAA
16451 AATTCACAGT ACAGAAGCAT ATGAACATAA TAAAAGTTCT TACCTCTTGT
16501 CTCCAGCAAG GAGTAAGTGT TTCAACCTGA AGGTTGGTTT TGAATTGTGT
16551 TCTGTGGAGC GTACTTAAAG TGAGTGAAGA AGAAAAATTT ATGTCAATCA
16601 TGATCATTGC AGCTGAAGTT TTTATTGTTT CACCCCTTAA AGGTTATTAA
16651 AATAGTATGT AGTTTAGTAG TCTTGATAAT TTTCCCTTAA GATTTATTGG
16701 CCAGTATATC AGGATTTTGT TTTAAATTTG ATATGTGAGC TTAGTTTTAT
16751 GCTATTTTCA AATAAGACAT TTAGAAGAAG ATAAAATAAC ATTCCTGTCT
16801 TAGTCTGTTT CTGTCTGCTA TAACAGAATA GCACAGACTG GGTAATTTAT
16851 AAACAGTAGA AGTTTATTTG GCCTGTGGTT CTGGAGGCTG GGAACCTCAA
16901 GAGCATGGTT CTGCCCTTTG TGCTGTGTTA TCATATGGTG GAAGGTGGAA
16951 AGGCAAGTGG GTATGTCAAG ACAGAGAGCA AGAAGGGGCT TGAACCTACT
17001 TTTATAACAG AGTGACTCCA GAGATAGCTA ACCCACTTTT GAGAGAATGC
17051 ATTAATCCAT TCATGAGGGC AGAGCCCTTG TGACCTAATC ACCTCTCATT
17101 AGGCTCTGCA TCCTTAAACT GGTTTTTTTT TGTTTTTTTT TTTTGAGACG
17151 GAGTCTCGCT CTGTTGCCCA GGCCGACTG CGGACTGCAG TGGCGCAATC
17201 TCGGCTCACT GCAAGCTCCG CCTCCCGGGT TCACGCCATT CTCCTGCCTC
17251 AGCCTCCCGA GTAGCTGGGA CTACAGGCGC CCGCCACCGT GCCCGGCTAA
17301 TTTTTTGTAT TTTTTTAGTA GAGACGGGGT TTCACCTTGT TAGCCAGGAT
17351 GGTCTCGATC TCCTGACCTC ATGATCCACC CGCCTCGGCC TCCCAAAGTG
17401 CTGGGATTAC AGGCGTGAGC CACCGCGCCC GGCCCCCTT AAACCTGTTGT
17451 ATTGGGGATT AAGTATCTAA CACAGGAAT TTGGAGGATA CATTTAAACC
17501 ATAAGAATTC ATGTATGCA AATGAATCCA TTCTAGATGA AAGAGAATGA
17551 ATTTAGTTTC CATTGAACCT TATAAATAGG CCTTTTCTAA GGTACTTACA
17601 GCTGATATTA TAAAATTTAT ATTTGTTTTT ATAAATTTGT ATTTGTATTT
17651 CTGTTTGTAC AAATACAATT ATACACTATA GTTCTCTGCT GTTAGATTTT
17701 TTTTCTTCCT TAGCATGTTT CCAAGGGGTG GAATGTTGAA AGTTGGGTTA
17751 ATGTCAATCA GCTTTCTTTT GTAAAGTGTT CATTGACATG TGAACCTTGT
17801 CTGAGAATCT AAATTTTATT TCATGAAAGA AGAAAACAGT ATATTCTCAT
17851 TTAACCCAGA ATTTAACTTC ATATACTTGT GGCTGTATTG GGAGTATGCC
17901 ATTGCTGTCT GTTTACAACC TGACCTACTC TACCTACTTA GAAGTAATTT
17951 GTGTTATGAT AGGTGTGCTG TGCTGACATA TGCTGAACAT ATTTGTAAGG
18001 GTGTTAAGTC ATTGAATAAA ACGCTTTTCT CCTCCTTTCA AATAACATTT
18051 TTTATTTCTG GTTATAAAAG TCATACAAGC TTAGTGCAGG TTGTTAAAAA
18101 GGTATAAAGA AGAAACCGTC AATCCATTAT AATCCTACAG TTTAGACTTC
18151 CTGCTCCAGC CTCTCAGAGT GCTGAGATGA GCTAGCCATG CCCAGCCCTC
18201 CAAAAGATTT TTTAAAAAAC AAAAATGAGG TTATACTTTA AAAAATCTTA
18251 TATTCCTTTC ACATAACAGT GTTATTTTGG AGGTTTTAGA ATTTCCAGTA
18301 GCATTTTAGA TTCAGAAAACA AGCTGATTCA TCCTCTACTT TGTACTTTAG
18351 GCAAGAAAAG AATTTTACCT AAATAGAATT TTGAACCTGAA AATCTGTTTT
18401 TCTAACTTTT TATTTAAAGA ATATTGTTCC ATGCTTTCAC AGTAGTGACT
18451 TTTAATTTT ATATTTTTAT TTTTATTTAT TTAGAGATGG GGGTCTCACT
18501 CTTGTGCTCT AGGCTAGAGT GAGTGCAATG GTTCTATTCC TAGCTCACTG
18551 CAACCTTGAA CTCCTGGGCT CAAGTTACCC TCCTGCCTCA GCCTTCTAAG
18601 TAGCTGGGAC TACAGGTGTG CACCACTGCA CCAGGCTTTT TTTAAAGGCA
18651 TAGAAAATGG TAGTGCTTGC ATACAAAAAT GGCGTAGGTA CATACATCAG
18701 CGGACATCAA GACTATGTTT AGATCATAAA TGTACATATA TGTACCGATG
18751 CCATTTTTC ACGCAACAA ATAATGGAAA TTGAACCTTA AACTGAAATT
18801 TGAAACAAGG GTTCTGGGGT GGGCCCTCTT GCTGATTGTT AATTGAATGT
18851 ATAGTTCAAT TTTTCCCAT CTGTTAAGCA AAAGACAATT CTAATGTTAG

FIGURE 3, page 6 of 21

18901 CAAAAATCCA CATATCCTGT CATTGATCAT TTTTTCCTTA ATTTTCTTTA
18951 AGAGATGGGG CTTCTCTCTA TGTGCCCAG GCTGGTCTGG AACTCTTGGG
19001 CTCAAATGAT CCTCCAGCCT CAGCCTCCCA AAGTGCTGGA ATTAATAGGC
19051 ACAAGCTGCT GTGCCTGGCC CTGTCATCAG TCATTTAACT TCATGCAAAC
19101 TGAGTAGAAT AAAACTCGTC CTTACTGTAC CTTATTGCTT TTGTTTTATT
19151 GTTGAACCT CCAATATTGC GAAAGTAGAC CAAAAGTTGA CTTATAGGAA
19201 AACTGATAG CAAAAATAAT TTTTCTCTTG TTGCTGTATT TCATGCCCAC
19251 CATCCAGTTG TTAAAGCCTA CTGTTAATTT CTCTCAGCCT CCTCCTTTCT
19301 GTCCAGGCTT ATTCTATGCC ATTCTTACCT TAACTGTTTT TAGCTTTCTC
19351 ATAGAGTGAA CTTTTTAAAT TAAAATAAAA TATCTGCTCG TAGTATTATA
19401 AAATTCAAGC AGTTCAACAG AATTTTTTAC TAATAGAAAT ACTTGACCT
19451 CAAAAGCAGC TTTATTTTAC AAACCCAGCC CAATTTGTGA TTAGATTTAA
19501 CTTGAGAAAA CATGAAATGT CTCTCATATT GTTTAAAAAT ATCATAAGTG
19551 GCTGGGCACG GTGGCTTATG CCTATAATCC CAACACTTTG GGAGGCTGAG
19601 GCAGTGGAT CACTTGAGGT CAGGAGTTTG AGACCAGCCA GGNNNNNNNN
19651 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19701 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19751 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19801 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19851 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19901 NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN
19951 NNNNNNNNNN NNNNNNNNNN NNNNNNNNTTC ACCATGTTGG CCAGGCTGGT
20001 CTCAAACCTC TGACCTCAGG TGATCCACCT GCCTGGGCCT CCCAAAGTGC
20051 TGGGATTATA GGCTTGAGCC TCGCCTGGCC TCCTCATAAT TTTTAAACCT
20101 TTATAAAAAC CTTTTCTAAA ACCCTTTTTA TTTTGAACCTA AATTTAGATT
20151 TACTGAAATT GTGAAATCAA TGTGGAGTTC TTGTATACCC TTCTTTCCGC
20201 TTTTCTTAAT AGTAACATCT TACATACATG GTACATTTGT CCAAATTAAG
20251 AAATAAACAT TGGTACAGTG TTAACATAG ACTTAATCTG GTTTCTCTAA
20301 TTTTCTTACT AATGTTCTTT TTCTGTTCTA GGATCTAATT CAGTATACCA
20351 TATTGTATTT AGTTGTAGGC CATGTTAGCC ACCTTCAATC TGTGACAGTT
20401 TCTCAGTCTT TCCTTCTTTT TCGTTATCTT GACAAGTTTG AAGAGTGCTG
20451 ATAGGTATTT TATAGAATGT CCGTCAGTTG TCTGTCAGTT TGTATTTGTC
20501 TGATGTATTT TTTTTTTTTT TTTTGAGATG GTGTCTCGCT CTGTCGCCTA
20551 GGCTGGAGTG CAATGGCATG ATCTTGGCTC AATGCAGCCT CCACCTCCGG
20601 GGTTCAGTG ACTGTCTGCG CTCAGTCTCC CAAGTAACTG AAACACAGG
20651 CATGTGCCAC CACGCCTGGC TAATTTTTTG TATTTTAGTA GAGAAGCAGT
20701 TTCACCGTGT TGCCAGGCT GGTCTCGTGC TCCTGAGCTC AGGCAATCCA
20751 CCCGCATTGG CCTCCCAAAG CGCTAGGATT ACAGGTGTGA GCCACCATGC
20801 CTGGCCAATA TTTTGAGGGA TATACTTTGG TGAGGTCTAG CAGATATCCT
20851 GTTCTCTCTT AGTTTATCG ATTAATTTAG CATTTATCCA GTAAATCTTC
20901 CTTGCAGCAA TTATTTTTTC TTTTCTTTT TTCCTTAATT TTTTTTTTAA
20951 GAGATGGGAT CTCACTCTGT TGCCCAAGTT GGAATGCAGT AGTGAGTTCA
21001 TAGCTCACTG CAGCCTCAAA CTCTGGGCT CAAGTGATCC TTCTGCCTCA
21051 GCCTCTCAAG TAGCTGGGAC TACAGGCATA GACCACCACA CCCAGCTAAT
21101 TAAAAAAAAT ATTTTLAGAG ATGGGGGTTT TGCTATGTTG CTCAGGCTGG
21151 TCTTGAACCT GCTGGCCTCA TGTGATCCTT CTACCTCAGC CTTACAAGTA
21201 GGTGGGAATT ACAGGTGTGA GCCACCACAC CCAGCATTCG AGCAATTATT
21251 AATGTAGTGC TACTGGTCAT TTTCTGTTTT TCTCATTTCT TCAGCATGTG
21301 TTATTGACTT GTCTCTTCCC TCCCATTAT AATCATTTAT ACTGCTATGA
21351 ATTCATGAGT ATTTATTTTG TGAGTTATAA TCTAATACGT ACTTAATTTA
21401 TTTTGTGCCT CAAATTGTTT TGGCTTGGCC ATTTTTTTTT TTTTTTTTTG
21451 AGACGGTCTC GCTCTGCTGC CCAGGCTGGA GTGCAGTAGC GCCATCTCTT
21501 CTCACGTCAA CCTCCACCTC CCGGGTTCAA GCGATTCTCC TGCCCTCAGC
21551 TCCTGAGTAG CTGGGACTAC AGGCGTGTGC CGCCACACCC GTCTAATTTT
21601 TTGTATTTT AGTAGAGACA GGGTTTACC ATGTTAGCCA GGATGGTCTC
21651 GATCTCCTGA CCTCGTGATC TGCCCGCCTC AGCCTCAAAA AGTGCTGGGA
21701 TTACAGGTGT GAGCCACCAA GCCCGACCG CTCTGTATC CTTTAAACAT
21751 GAGGIGCTGT CATCATTTTT TCCCCCTAAT ATTTTGGCCA AAAATGTTAA
21801 TCAAGGATGG CACAAATTTT CTGTAGCTGT ATCTCACAAT GAAAGAGGCC
21851 TGATTAAAAA TGTA AAAACTA AAATGTTCTC TGATCTCTTA GCACATGCTT
21901 TGTAAAAGGC ACAGTGCTAG ATCCTTGAT ACGTAGATGA GTAAGTCAGC
21951 TTACCTTCCA CAGCCACAGA TAGCTATGTC AAACGTAAGG GTGGAGAAAC
22001 ACAGACCCCA AACTTCTCGA GGGTAGAAAA TATGAGGTTA TAGTAGATTA

FIGURE 3, page 7 of 21

5'-TTCAGTCCCC-3' 5'-TCTTCTGATT-3' 5'-TACTTGTTTA-3' 5'-GAAGATTTTT-3' 5'-GTTTCCTTCT-3'
5'-CTGACTTCTA-3' 5'-TTTTGCTGCT-3' 5'-GACTGGCACT-3' 5'-TGGGATTTTT-3' 5'-AAAAAATTAT-3'
5'-TTTCCTCATA-3' 5'-TATAATTTAA-3' 5'-GACAATAAGT-3' 5'-ATAACAATAA-3' 5'-GTATAATATG-3'
5'-GTAATTTGCT-3' 5'-AAAACCCAAA-3' 5'-CAATGTTTTA-3' 5'-AGTAATGCAT-3' 5'-ATCATTATGT-3'
5'-AAACCTACGT-3' 5'-AATAGTTGAA-3' 5'-TATTCACAAA-3' 5'-GATAATCGCT-3' 5'-TATAGAAGTT-3'
5'-TTATATCCTC-3' 5'-TCTTCTTTGG-3' 5'-CAGTGCAATT-3' 5'-AAAACAAAAA-3' 5'-AAATAAGTTT-3'
5'-TATGTCTTGT-3' 5'-TTACATGTAA-3' 5'-ATAATTTTAA-3' 5'-TCTAAATTGT-3' 5'-GACGTGGTTT-3'
5'-TCACTTTAGC-3' 5'-ATATTTTTGA-3' 5'-AAGTAAATCA-3' 5'-AAAAGGACAA-3' 5'-AATACAAAAT-3'
5'-CATGTATATC-3' 5'-TTCTACAAAA-3' 5'-ACGATATATA-3' 5'-AATTCTAAGG-3' 5'-TTTTTGTCCT-3'
5'-TTTGAAATTG-3' 5'-CTTAAAAGAA-3' 5'-TGCATAGAAC-3' 5'-TGGTGTCTGA-3' 5'-GTTGGGAAGG-3'
5'-ATCTATGAGG-3' 5'-GATTTTCCTT-3' 5'-GAGACCGTGG-3' 5'-GTGAATAATA-3' 5'-ATGTTGTCTT-3'
5'-AGTTCCATGA-3' 5'-AGGAATCTCT-3' 5'-GGGGATAGTT-3' 5'-TTTGAGTTAG-3' 5'-GCCTGGCAAT-3'
5'-GTTAGAGATA-3' 5'-CATAAAGAGA-3' 5'-GCCTTGTTTT-3' 5'-ATCACTGGGT-3' 5'-GCGGTGGCTC-3'
5'-ACACCTGTAA-3' 5'-TTCCAGCACT-3' 5'-TTGGGAGGCT-3' 5'-GAGGCGGGCA-3' 5'-GATCATGAGG-3'
5'-TCAGGAGATC-3' 5'-GAGACCATCC-3' 5'-TGGCCAACAC-3' 5'-GGTGAAACCC-3' 5'-GTGTCTACTA-3'
5'-AAAATACAAA-3' 5'-AATTAGCTGG-3' 5'-CGTGCTGGC-3' 5'-GCATGCCTAT-3' 5'-AATCCCAGCT-3'
5'-ACTCGGGAGG-3' 5'-CTGAGGCAGG-3' 5'-AGAATCACTT-3' 5'-GAACCAAGGA-3' 5'-GTTGGAGGTT-3'
5'-GCAGTGAGCC-3' 5'-GAGATCGCGC-3' 5'-CACTGCACTC-3' 5'-CAGCCTGGGT-3' 5'-GACAGAGCAA-3'
5'-GACTCCGTCT-3' 5'-CAAAAAAAA-3' 5'-AAGCTTGGTT-3' 5'-TTCAATGGTT-3' 5'-CTGAAAAATG-3'
5'-CTTTAATACA-3' 5'-AGTGTAGAGT-3' 5'-GTTAGTCAAG-3' 5'-TTTTGCACTT-3' 5'-GGATAAACAG-3'
5'-CCTGTGAATT-3' 5'-TATCACATTT-3' 5'-CTAGTTTATA-3' 5'-ATATGGGCTT-3' 5'-TCAGAAGTTA-3'
5'-TATGAACATT-3' 5'-GTTTTGACGG-3' 5'-GAGAATTCAT-3' 5'-GCTGGATGCT-3' 5'-AGAGAAGGAT-3'
5'-CGTGAGAACC-3' 5'-CCTTCATTGG-3' 5'-AGGAGTGCTA-3' 5'-TGAAATTATT-3' 5'-TGATCTTGGA-3'
5'-ATTTTTTTTT-3' 5'-TTTTTTTTTT-3' 5'-TTTTTTTTTT-3' 5'-TTTTTGAGAC-3' 5'-AGAGTTTCGT-3'
5'-TCTTATTGCC-3' 5'-CAGGCTGGAG-3' 5'-CTGGAATGCA-3' 5'-GTGGCACGAT-3' 5'-CTCGGCTCAC-3'
5'-TGCAACCTCT-3' 5'-GCCTCCTGGG-3' 5'-TTCAAGCAAT-3' 5'-TCTTCTGCCT-3' 5'-CAGCCTACCA-3'
5'-GGTAGCTGGG-3' 5'-ATTACAGGCA-3' 5'-TGCGCAACCA-3' 5'-TGCCCAGCTA-3' 5'-ATTTTTGTAT-3'
5'-TTTTAATGGA-3' 5'-GACGGGGTTT-3' 5'-CACCATGTTG-3' 5'-GTCAGGCTGG-3' 5'-TCTTGAATCT-3'
5'-CTGACCTCAA-3' 5'-GTGAATGCC-3' 5'-TGCCTCAGCC-3' 5'-TCCCAAAGTG-3' 5'-TTGGGATTAC-3'
5'-AGGTGTGAGC-3' 5'-CACTCGCCCT-3' 5'-GGCCTGATCT-3' 5'-TAGAATTTGA-3' 5'-AGGAGAGACT-3'
5'-AATATTTTAT-3' 5'-GGGCAAAAAC-3' 5'-AATGAAAAGT-3' 5'-TACCTTTCTG-3' 5'-TATTCTAATA-3'
5'-CTATAGAGGA-3' 5'-GTGGGATTTA-3' 5'-TTTAGAATGT-3' 5'-TTTAAGTATC-3' 5'-TTGGGCAGTC-3'
5'-CAAGAGTGCG-3' 5'-TATCACTTAT-3' 5'-TTTTCTTTTC-3' 5'-CTTCTTTCTT-3' 5'-TTTAAGTGGA-3'
5'-AGTTCACTGA-3' 5'-TGTTAGAGAT-3' 5'-CATAGGTGGC-3' 5'-ATTGCCTACT-3' 5'-TTTTACATAA-3'
5'-TTTTATCATG-3' 5'-TTTAGTGATC-3' 5'-TGTGCAAGG-3' 5'-GCTGTGGCTG-3' 5'-TTTGCAGTTT-3'
5'-TGGCTTAAGC-3' 5'-TGGCTTAAGC-3' 5'-CTTTATAGGA-3' 5'-GATGTAGTCT-3' 5'-TCACAGTGAG-3'
5'-TTGTTATTTG-3' 5'-TAGCTGTGTT-3' 5'-TTTGTTTTTG-3' 5'-TATAGCTTAT-3' 5'-AGCAATGCAG-3'
5'-TGTGCTTTTT-3' 5'-ATTAACATCA-3' 5'-TTTTCTTTTT-3' 5'-CTTTTTCAG-3' 5'-TGATTATTTA-3'
5'-TTCAAGTTAC-3' 5'-TTCTGATTGG-3' 5'-CGACTCAGGG-3' 5'-GTTGGAAAGT-3' 5'-CTTGCCTTCT-3'
5'-TCTTAGGTTT-3' 5'-GCAGTAAGTT-3' 5'-GAAATTGAAA-3' 5'-TGTCTTTACA-3' 5'-ATTAATGGTA-3'
5'-CAATTAATGC-3' 5'-TATGTATGTT-3' 5'-TTCTAGGTAG-3' 5'-ATAAAATTAA-3' 5'-ACAGTTTATAT-3'
5'-TCAGAATAAG-3' 5'-TTAATTCTTC-3' 5'-CAGAATTTAT-3' 5'-ATATTTAAAG-3' 5'-ACTCCAAATA-3'
5'-TACATCCCCA-3' 5'-GTGGTATCTT-3' 5'-GGACTGTAA-3' 5'-ATAGAAAAAT-3' 5'-ATTGTTGCTC-3'
5'-TTAAAAGAAA-3' 5'-TTCAGTGAAG-3' 5'-TCTGGTTATA-3' 5'-AAGTCAGAAT-3' 5'-GTCTAATACT-3'
5'-TTTGGTCAGA-3' 5'-GTCAAACAGC-3' 5'-AGTTCCAATA-3' 5'-TAGGCAGCAA-3' 5'-GTTAAAGGGG-3'
5'-TAGTTGGTGG-3' 5'-CCTGTGTTGA-3' 5'-AAGCGACTTG-3' 5'-ATGAAAATAA-3' 5'-ATCTTTAAAT-3'
5'-TAAACTTTAG-3' 5'-TAGAATAAAA-3' 5'-AGAAAAAGCA-3' 5'-GAGCCAGGTG-3' 5'-ACGCAGTGGA-3'
5'-TCATGCCCTG-3' 5'-AGTCTCAGCT-3' 5'-ACTCAGGGTG-3' 5'-CTGAGGGTGG-3' 5'-AAGGATCACT-3'
5'-TGAGTCTAGG-3' 5'-AGTTTTGAGA-3' 5'-CCAACCTGGA-3' 5'-CAACATAGCA-3' 5'-TGACTCTGTC-3'
5'-TCTGAAAAAA-3' 5'-AAAGTTAATA-3' 5'-AAAGAAAAAG-3' 5'-TAGGGTCTTG-3' 5'-GACAAACTTC-3'
5'-GTTGGCCAAT-3' 5'-GGCATAGTTC-3' 5'-TAAATGCTGA-3' 5'-AGCTGACAGA-3' 5'-TAAAGGACTT-3'
5'-TTGACTTAAC-3' 5'-AGAATCCACA-3' 5'-GTGTCCTTCA-3' 5'-TAGTCTTTAT-3' 5'-CAACTACCTT-3'
5'-TAAATTTAGC-3' 5'-ATGTTTCCTG-3' 5'-GCCAGGTGCG-3' 5'-GTGGCTCACG-3' 5'-CCTGTAATCC-3'
5'-CAGCACTTTG-3' 5'-GGAGGCCGAG-3' 5'-ACGGGCGGAT-3' 5'-CACAAGGTCA-3' 5'-AGAGATTGAG-3'
5'-ACCATCCTGG-3' 5'-CTAACACGGT-3' 5'-GAAACCCCGT-3' 5'-CTCTACTAAA-3' 5'-AATACAAAAA-3'
5'-ATCAGCTGGG-3' 5'-TGTGTTGCCA-3' 5'-CACGCCTGTA-3' 5'-GTCCCAGCTA-3' 5'-CTCGGGAGGC-3'
5'-TGAGGCAGGA-3' 5'-GAATCGCTTG-3' 5'-AAGCCAGGAG-3' 5'-GCGGAGGTTG-3' 5'-CAGTGAGCTG-3'
5'-AGATGGTGCC-3' 5'-ACTGCACTCC-3' 5'-AGCCTGGCAA-3' 5'-CAGAGCAAGA-3' 5'-CTGTCTCAAA-3'
5'-AAAAAAAGAA-3' 5'-AAAAAATAAA-3' 5'-AAAACAAATT-3' 5'-AGCATGTTTC-3' 5'-CCTTCTAGAG-3'
5'-ATCATTGTTT-3' 5'-CTCAGAGCAT-3' 5'-GGACCAAGA-3' 5'-CTCCTGGGGG-3' 5'-TTACCAAGAC-3'
5'-CCTCTCAGGT-3' 5'-AGCCCATGAG-3' 5'-GTCAAAATAT-3' 5'-CCTAATAATA-3' 5'-CTAAGATGTT-3'
5'-AGTATTTGTA-3' 5'-AGGAAATATT-3' 5'-TACTTGGTAA-3' 5'-TAATACTAAT-3' 5'-ATAAAGATG-3'
5'-TTTGCCTTTT-3' 5'-TCAGTGATGA-3' 5'-CATTGGCTCT-3' 5'-GGTACAAAAG-3' 5'-CATGTGGGTA-3'

FIGURE 3, page 9 of 21


```

31501 TTTTTTTTTT TTTTGAGACA GAGACATGCC TTGTCGCCTG GGCTGGAGTG
31551 CAGTGGCACA GTCTCGGCTC ACTGCAACCT TCACCTCCCG GGTTCAAGTG
31601 ATTCTCCTTC CTCAGCCTCC CGGGTAGCTG GGATTACAGG CGCCCGCCAC
31651 CACGTCCAGC TAATTTTTTG TATTTTAAAT AGAGATGAGG TTTCACCATC
31701 TTGGCCAGGC TGATCTCAAA CTCCTGACCT TGTGATCCAC TCACCTGGGC
31751 CTCCCAAAGT GCTGGTATTA CAGGTGTGAG CCATCATGCC CGGCCCATAT
31801 TTCTAAAAAC ATTTTCTTAT AAAATGACAT TGCCATTATC AACCTGCAAA
31851 ATACATTCC ATTTGGTTGT TTTCTTGCTT AGTCTTTTAA TCTAGAGTTT
31901 TATACCTTAT CTTTTTTATT TATATATTTT TTATGTCATT GACTTTTTTG
31951 AGAAACTGAA GCACTTGCTC TGTAGATTGT CCAATATTCT AGATTGTGCA
32001 TTTTGTTTCC TTGTGATGTC CTTATGCTTA TTTGTTTGTC CCTCTTTCTG
32051 TAATTAGAAG ACCTAGAACT GCACTATCCT TAGAGTAGCT ACTAGCTCTA
32101 TGTAGCTATT TAAATTTAAA TTAATTAATA TTGAAAAAGT TTGGTGGCTC
32151 ACACCTGTAA TCCCAGCACT TTGGGAGGCC AAGGTGGGAG GATTGCTTGA
32201 GTGCAGGAGT TCAAGGCTTC AGTAAGCTAC GATTGTACTC TAGCCTGGGA
32251 GACATCAAGA CCCTGTCCCT TTAAGGGGGA AAAATAATTG AAAAAATCAA
32301 AAACCTTAGT TCCTTGTTTC ACAAGCTGCA TAGGGCTAAT GGCTACCATA
32351 TTGGCTAGCA CAGCTTATAG AACCTTTCCA TTGTCACAGA AAGTTCTGTT
32401 TGGCAGTGCC GTTCTCATTA GACCTGATTC GATTAAGGTC CATCTTTGTT
32451 GACAGAGTAC TTCTTAGGTG GTGCTTTGTG GTTCATATGA TGATAGCCTG
32501 GTCTGTTTAT TCATATATCT TTTACAGAGA AATATTTTAA TTCCATTCTG
32551 AATAAAATTT CATGGCAGGT ACTTGCAAGA AGCAGTTATA ATTTTAAAGT
32601 TTAACATTAG GTTAAAAAAT TGACAGGAAA CATATATTCA CAGGTAAAAAC
32651 TTGTACACAA ATGTTTCATGG CAGCATTATT CATAATAGCC AAGAAGTGGA
32701 AACAAACCAA ATCAATTTAT GAATGGATAA AATGTTGTAT ATTTGTAGTA
32751 CATGTAATAT TATTCAGCCA ATAAAATGGG CCAGGCATGG TGGCTCACAC
32801 CTGTAATCCC AGCACTTTGA GAGGCTCAGG CAGGGGGATC ACTAGAGGTC
32851 AGGAGTTTGA GACCAGCCTG ACCATCATCA CGAAACCCCTG TCTCTACTAA
32901 ACGTACAAAA ATTAGGCAGG CGTGGTGATG CACGCCTGTA GTCCCTACTA
32951 CTCAGGTGGC TGAGTCATGA GGATTGCTTG GACCCCGGGA GACAGAGGTT
33001 GCAGTGAGCT GAGATCATGA CACTGCACTC CAGCATGGGC AACAGAGCAA
33051 CATCCTGCCT CAAAAAATAA AAAAAAATAA AAAAGAAGTA CTGTTACATG
33101 GTACAACATG GATGAACCTT GAAAACATTC TGCTAAATGA AGGAAGACAG
33151 ACACAGAGGG CCACATATTT TATGATTCCA TTTATACGAA ATGTCCAAAA
33201 TTGGCAAATC TAAAGAGAAA GTAGATTAGT GGTGCGCAGG GAGTGAAGAC
33251 GGGTTCTTTT TGGAGTGAAG AAAATGTCCT GGAATTCGTG GTTGTAGTTT
33301 GCAACCTTGT GAATGCATAA GGACCACTGA ATTGTCCACT TCAAAAGGGT
33351 GACTTTTATG TTATGTGCAT TATATCTAAA AAAAAATCA TAATTAGGAA
33401 GCAAGATTGA CTTCTAAGAA AAAGCGGAGT GAAATTGTTG TTTTGTGGTG
33451 AATAAATTGG GTGGGTGGGT CGCAAGAGTT TTGCTGATTA GTGATTAGAA
33501 AAATTATTCA TAATCATTGA AAATATAAAA TATTTTCTTA TATGATGTAT
33551 GTAAAGAATT TGGCAAGAGA TGATGTTTGG AAAAAATAAA GAATGGCTAT
33601 TGTAGAGATC TTAAGGAAAG AAACCTACAGT TAAGTAGTGC TTTGTAATCA
33651 GAATATGAAG TAAGTACTGA AAGTGGATGG AGTGGCTGTT GTCAGCATGT
33701 TATACTTTAT ACATTTTATT CATAAATTG GACTGTAGAT AAAAGTAAAC
33751 TTTTTTTTTT TTTACTCTTG AACAACAGTT TTTTTTTTTT CACTTAGACT
33801 TGCATCTGCT CCACTGAACA ATACATTTAA TTGTTAATTA TTTCCCCCTT
33851 CAGGATGATA CATATACAGA AAGCTACATC AGCACAATTG GTGTGGATTT
33901 CAAAATAAGA ACTATAGAGT TAGACGGGAA AACAATCAAG CTTCAAATAG
33951 TAAGTGACTT GGCTAGTAAT TTTTTTGAAA TTTATTTTGG TAAATTTGTA
34001 ATGTATTGTT ATTTTGTATA TATTTACTAT GCTAACAAAA TTGAATGTAA
34051 AATGTCTTAA GATTCATGTA CTTAAGATAG AATGGTAGAA TAAGAATTAC
34101 TTAGATTAAA AATAATATTT TCAAGATTAC TTAAGCCTCA TTGAATTTTC
34151 TGTTCATGAA GCAGAGAAAC TCATGTTTTA AGTCAAACCT GGTCCCTCATC
34201 TTTTCTTTT ATCAGTGGAA ATCTAAGTTC AAGTTTACCT TGTCTACAC
34251 TGCAATGTT ATAGACCAT TTTGTTTGTC TTTTACTGTG CTAAGTGCAT
34301 GGAACATTAA AGGAACCCTA GGAAGAGATT CTTCATATGT GGCTCAGTTG
34351 AAGAGAAGTA CTTATGTAGT TCTAAGTATT TTTATTAGAT AGTGTGCACC
34401 AACTCTGTAG AAACACAGAA TTTTGTGGA AAAAGGAACT TAGTTTTTGT
34451 AACATGTTCA TTTTACTGCT CAAAAAACG AATGCTGAAA GATTTAATGA
34501 CTTGCCTACA GTTACTGGTA GAACCAAGTG ACCGAAGCTC TGTCTTCAAT
34551 ATTTTGTGTC TGTGTGCCAT CCTATCCCCC TTATCCATCT TTACACCCCC
34601 AGCCCCCAAT TAAATATAGG CAATTATAAT AGTTCAGTTG TGCTCTTCA

```

FIGURE 3, page 11 of 21

34651 GTATGGGTCT GAGTCCCTGTC AGTGTGGGCA TATCTGTGGT CTTTTAAAAA
34701 ATAAATCTCT CAGTATTTT CAGAGTAGGC TATTAGCAAG AAGTAGGCTA
34751 TAAACACAGG AAACCGAGTGA CTGCCCCCTT TCATGGAAGT GATGACACAT
34801 GGAATTGGAA GGAGTCTGCTC ATTAGGAGTC AGAAGACTTA GATTTGTGTG
34851 CTTGGTTCTA GTATTTACCT GTTAGAGAAT CATGGGTTTG TGTCTCTGGG
34901 GAAAAGGCCG AAGTAACCCCT GAGACCCAGT TTCCTTTCTA AAATGTGTGT
34951 GATGACACCT GATTTACTAA TTTATAAGCT AGTTGTGAGA ACCAACTGTA
35001 ATAGCTTTGT GTATGTGACA ATACGTGTGA AAGCCCTTTG TAAACTTTTG
35051 GGCAGCATAT AGATACTACT TATGATATGA CATGCCAGG TAAATGGGTG
35101 TTTGATAGGT TAAGTTGCTC CCTTTTCTTA CATGACTCTG ATGAGGAAAA
35151 GAAGGTATGT TAACAAAAGA TAGGTGGCTG TGGATATTGA TATAAGTAAA
35201 CACACTTGAT GTGTCAAATT AGGACTTGCA AGGATTTAGT TTTAGAAAAT
35251 AGCTTGAAAT ACTTTCAATC AGTGAACAAA TTACCCTCCA TATTTTTTCC
35301 CACGATATAA GTACAGTCTC AACCTTTTAT TTGGCACCAT AAAGAGCACA
35351 TAAAGATCTA CCCAAACTG TACTTTAAAG CACTGGTATG GAATAATTGT
35401 ATTATGTGTG ATCATTTGGT TTTATAAGAT TTGGGTGTGT ATTCTGTGTG
35451 GAAACATTCA TATTTTGTTA CTTTCTGTG GCTGGAAGG ATCTTATAGG
35501 ACACTGTCTT TCATCTTTGT CTGTCTTTCA TCTTTAATAG GAATTTCTTT
35551 TCCATGCCTG AAGGCCTCAT TTTGAACATT TTGTTTGTG GTTTTTTTAT
35601 TTTTTGAGAT ACAGTATTGC TCTGTCTCCC AGGCTGGAGT GCAGTGGCGC
35651 GATTTGAGCT CACTGCAACC TCCGCCTCCT GGGTTCAAGT GATTCTCCTG
35701 CCTCAGCCTC CCTAATAGCT GGGATTACAT GTGTGTACCA CCATGCCCGG
35751 ACAATTTTTT TTTTPTGAG ATGGAGCCTT GCTTTGTGCG CCAGGCTGGA
35801 GTGCCAGTGG TGCAATCTTG GCTCGCTGCA GCCTCCGCCT CCCAGGTTCA
35851 AGCAGTTCTC TTGCCTCAGC CTCCTGAGTA GCTGGGATTA CAGGCGTGCG
35901 CCACCACACC CTGCTAATTT TTTGTATTTT TAGTAGAGAC AGAGTTTCAC
35951 CATGTTGGTT AGGCTGGTCT CGAACTCCTG ACCTCGTGAT CTGCCTGACT
36001 CGGCTTCCCA AAGTGCTGGG ATTACAGGCA TGAGCCACTG TGCCCAGCCT
36051 TCCGATAATT TTTGTATTTT TCGTAGAGAT GGGATTTCGC CATGTTGGCC
36101 AGGCTGGTCT CAAACTCCTT ACCTCAAGTG ATCCACCCGT CTTGGCCTCC
36151 CAAAGTGCTG GGATTACAGG CGTGAGCCAC CACGCCTGGG TTTTTGAACA
36201 TTTTTAAGAA GCTTACCATT TTTTCGAAAT AGCTAGTTCC ATTTTACACA
36251 TAACTTCAGC TAGGCATGTT GCCTCATGCC TGTAATCCCA GCACTTTGGG
36301 AGGCCGAGGT CAGAGAGTCA CTTGAGGCCA GGAGTCAACA TAGCTCCTGT
36351 GACCAGCCTG GTCACATAG AGACTCTATC TCTACCAAAA AAAAAAAAAA
36401 AAAAAAGTAAC CAGGTGTGGT GGTCCATGCC TGTAGTCTTA GCTCCCCAGG
36451 AGACTGAGGT GGGAGGAATG TTTGAGCCCA GGACTTCAAG GCTGCAGTGA
36501 GGCAAGATTG CACCATTGCA CCCAGCTTTT GGGGACAGAG TGAGAGACCC
36551 TGTCTCAAAA ACAAATAAG GCTGGGCGCA GTGGCTGTCC GGGCGTCTG
36601 GTTCACGCTT ATAGTCTTAG CACTTTGGGA GGCCAAGGTG GGCAGATTGC
36651 CTGAGCTCAG GAGGTCTAAG ACCAGCCTGA GCAACATGGC GAAACCTCAT
36701 CTTTGCAAAA CATAAGAAA AAAACAAAAA AAACCACAAA ACCTCTAGTT
36751 GCCAGTTATT TTTTCTATTT ATTCCTAGTG ATTCTTCTTT TTTTCTTTT
36801 TCTGAGACAA AAATTTCACT TTGTCTCCCT CGCTAGAGTG CAGCGTCTAG
36851 CTCCTACAT GATTCTTTTA GAGACATGTT AATTCTTTAT ATTGAGCTGA
36901 AGCCTGTTTC TTTTACTTCT GTCTCTTCTT ATTCCTCCGC CTTGTAGAGC
36951 TGCCTGAATC AGATTAATTC CTCTTTTATT GGCAAGCCTG CCCTTCAGAT
37001 TGATCTTATC ACAACCTTTC TTCTACCTCT GAAGTCCTCA TTCTTTCCTG
37051 TAATGATATT TTCAGAACCT TGTGCAATTT GGGTTATTCT TACATTTTAT
37101 AAATGCCTTT TATTAATTTT GATTTCTTAA ATCAAGTATG AGATATAACA
37151 CATGAGGTAA ATCCTGTCTT GATTTGGAGC CTGAATGAAT TTCTCTCTTG
37201 AACTTCAAGG GCTCATGGCC CTTTCTTATT ATTAATCAAA GACAACCATT
37251 TGTGTTTCA GTAGCTATAT TATTTCTAGT TTGGGTCTTA AGGTTTTTGA
37301 TTTGCTTGT TTTTCTTTTT TCTTTTTTTT TTTTTTGAGA CGGAGTTTCG
37351 CTCTTGTGTC CCAGACTGGG AGTGCAATGG CGTGATCTCG GCTCACTGCA
37401 ACCTCCGCCT CCCAGGTTCA AGCGATTCTT CTGCCTCAGC CTCCCTAGTA
37451 GCAGGGATTA CAGGCATGTG CCACCACGCC GGGCTAATTT TGTATTTTAA
37501 GTAGAGATGG GGTCTCTCCA TGTGGTCTAC GCTGGTCTCG AACTCCCGAC
37551 CTCAGGTGAT CCGCTGCCT TGGCCTCCCA AAGTGCTGGG ATTACAGTCG
37601 TGAGCCACGG CGCTTGCCCG ATTTGCTTGT TTTTAATTAA AATAGGGGCC
37651 TTGGCCAGGT GCAGTTGTTT ACCCCTGTAA TCCCAGTACT TTGGGAGGCT
37701 GAGGCAGGCA GATCTCTTGA GTTCAGGAGT TCAAGACCAG TATGGGCAAC
37751 ATGGTGAAAC CCTGTCTCTA CAAAAACAC AAAATTCAGC CAGGCATGGT

FIGURE 3, page 12 of 21

37801 GGTGTGTCCC TGTAAGTCAA GGTACTCAGG AGGCTGAGGT GGGAGGATTG
37851 CTTGAGCCCG GAGATGGAGG TTGCGGTGAG CCAAGATTGT GCCATTGCA
37901 CTCTAGCCTG GGCAACAGAG CGAGACCTTG TTCAAAAAA AAAAAAGAAG
37951 AGGGTCTCAC TTTACACTTC TGTGACTGGT GTTTTAAAAA TCTAAACACA
38001 GGCCGGGCAC GGTGGCTCAC GCCTGTAATC CCAGCACTTT GGGAGGCAGA
38051 GGCACGCAGA TCACAAGGTC AGGAGTTCGT GACCAGCCTG GCCAGCATGG
38101 TGAAGCCCAT CTCTACTAAA AATACAAAAA AATTAGCTGG GCATGGTGGC
38151 AGGTGCCTGT AATCCCAGCT ACTTGGGAGG CTGAGACAGG GGAATCACTT
38201 GAACCCAGGA GCGGAGATT GCAGTGAGCC AAGATTGCGC CATTGCACTC
38251 CAGCTTGTG ACAGAGCGAG ACTCCGTCTG AAAAAAAAAA AAAAAATCT
38301 AAACACAAGA TTTTACTTTT AATCCTATCA TTCTCTCTG CTTGGCTTCA
38351 GTAATCCTTC AAGTTTTCTA GGTCTTTTCA AAATCTTGAT TCTGTTGATT
38401 TATATTTTAA TTATCTTTTC CTTTCAGCTT TTCCTGTTCA GGTGTGACAT
38451 CTGGGTCTTT ATCTGAGTTT TATTAGATTA TAAACATTC AGCAAGATAG
38501 GGCAGGTACT GAGTCCAGTT GTACACCATG GAAGGCCTCT TTCTGTGATT
38551 GTTCATTAT GAGGCTTTAT GAAAATGTCT ACATTACACC AGGCACTTGG
38601 AGGTTACAGA GATGAATAAA ACATAGTCCA TTAGGAGGCA GACAATGGGA
38651 GAGACAAACA TGGGAAAAAG TTACTCTGAT TATGAGGAGT AATGAGAATT
38701 ACATATGAAG GAAAGTATTG TTAGTACTGT TAGGATTTAG TGTGAGAAA
38751 GTTTTCAGAG TAGCAAGGAA ACATCAGAAA TTTTACTCTT TCTGCCAGGC
38801 ATGGTGCATG TATTATTCTG TTCTCACACT GCCACAAGGA ACTGACCAA
38851 ACTGGGTGAT TTATTAATAA AAAGGTTTAA TTGACTCATA GTTCTGCATG
38901 GCTGAGGAGG CTTACGAAA CTTACTGTGG CAGAAGGGA AGCAGGCACG
38951 TCTTACATGG CAGGAGGCGA GAGAGTGTGA AGGAAGTGAA GGGGGAAGAG
39001 CCCCTTATGA GACCATCAGA TCTTGTGAGA ATTCATTAC TATCACTCGA
39051 ATGGGGGAAA CCGTCTCAT AATCCAATCA CTTCTCCATA ATCCAATCAC
39101 TTCCCTCAGT GATTACAAC TGAATGAGA TTTGGGTGGG GACACAGAGC
39151 CAAACCATAT CAGTGCTGT AGTCCCAGTT ACTTGGAGGC TGAGGCAGGA
39201 GGAACACTTG AGCCAGGAG TTCAAGATCT GCCTGGGCAA CATAGCAATA
39251 CCTCCATTTT GGTAAAAAAG GAAATTTTAC TTTTGGGTG CCATTGCTTA
39301 GTTTAATCAG CTGTAACCTC TTGTTGACTT TTAGTCAAAA AACAATTTT
39351 CCTTCTATCT TTGTGAAAGA GGTGGTGAG CAAGGAAGAA AAGGAACTT
39401 GCTTTATGTA GCAGTCTCTA TAGTCAGGCA CATTTTACAA ACATTAGTTC
39451 ATTTAAACCC CTTTAGCTGT TGTACAAGGT GAATGCTATC TAGCATTTAC
39501 AGATGAAGAA ACTGTTAGGT GACTCTCCCT AATATTAAAT AACCAGGAAC
39551 CTGGATTGTA TGTTTTGAAG TCAGGGTAGC TTGATCCTCG AGTTCATGCT
39601 TCCTCCAGG ATACACTGAA AGACTTTGAG CCTCTTTTTT TTTTTTCTC
39651 TTTTTTTGAG ACAGGATCTG GCTCTCTTGC CCAGAGTGCA GTGGTGTGAT
39701 CTCAGCTCAC TGCAACCTCT GCCTCCTGGG CTCAAGCGAT TCTGCCTCAG
39751 CCTCTCGAGT AGCTGGGACC ACAGGCGCAC GCCAGCATAC TTGGCTAATT
39801 TTTGGATTTT TAGTAGAGAC AGGGTTTCAC CATGTTGGTC AGGCTGGTCT
39851 CGAATCCTG AGCTCGTAAT CCGCCCGTCT CGGCCCCACA AAGTGTGGG
39901 ATTACAGCG GTAGCCACCG ACCAGTCCC AACAGTTTTT TAAAACCCAG
39951 AACTATAATG CAATAATGTT AGCATTTGTT TTGGGAGTTT GAGCCTAAAT
40001 GGTGGAAGTG CAGTAAATTG TTCTTAAAT ACGTTTTATG AAAGTATTTG
40051 GAGTCTCTTC CTTACATTTT TTTCTCTAGC ATGAAGACAA CACCTAGCCA
40101 GGCATGGTGG CTCATGCCAG TAATGCCAGC ACTTTGGGAG AATGAGTTAG
40151 GATAATTGCT TGAGTCCAGG AATTTGAGAC CAGCCTGGGC AATGTAGCGA
40201 GACTCTGTCT CTACAAAAA GAAAAAATTA GCCGGGTGTG GTGGCATGTG
40251 CCTGTAGTCC CAGACTCTCA GGAGGCTCAG GTGGAAGGAT TGCTTGAGGT
40301 GGGAGGTTGA GGCTGCAGCG AGCCATGATC ATGCCACTGT ACTCAGCCTG
40351 GATGACAGAA TGAGACGCTG CTTGAGAGGG GAAAAAAG ACACCTGCTT
40401 GGGATGATTA AAGTCTGTG TTAGCTGGTA GTTATTTGAA TTAGGTCCCT
40451 CCAGTGCTTT TAATCATGGT AGAATGTGCT AGCAAGTGAG TTTGTCTTAC
40501 ATGGAAGAGT TCTGTGTTCA AGGGCTTTTC GCCAGTGGCA TTCCTAAACA
40551 CAGTGTTAAA GCGGTAGGG AATGTGAAAA GTATGACATA GTTCTGCTC
40601 TCAACAGCTT GTAATTTTAG TATTATTATC GTAAGCTCAA TTGTAGGTAC
40651 TACTTCTTTT CTGGACTTTC AGGTGCTTAT TACCGTGCAA TTTAGTGGTA
40701 TGAGTTGAGG ACTAATGTTT CTATATCACA TCCTGATAAT CTCCACAGTT
40751 ATGAAACTA AACTATTTCC CCTCCCTCCT ACACTTTTCC CCAACTTTAT
40801 TTTAATGGAA TTGTTTGGAT TTCTTGATTG TTTTGTAAATA GTGGGACACA
40851 GCAGGCCAGG AAAGATTTCG AACAATCACC TCCAGTTATT ACAGAGGAGC
40901 CCATGGCATC ATAGTTGTGT ATGATGTGAC AGATCAGGTA AGTTCCAAGA

FIGURE 3, page 13 of 21

40951 GGAGATTGTG TTACAGTGAC CAAGTAGGAA GCCATTATTT GATTAATGTC
41001 AGATTCATTT ACTACTTCAT ATATAAGCCA TCAGTATTAA TTTTATGGCA
41051 GAAAACCTTG TCCACTCTCA AATATAAATG TGAATCACTT AAAAGACATT
41101 TGTTCCTCTG TAATAAATAA AAGATTAGTA ATTAGTTTTA CGTTTGCTTT
41151 CAAGGGATTC TGGTTGTATT TATTGTCAAC TAAATAACTT TGATCAAATA
41201 GCCAAGACTC TAACATATAG GCAAGAGTTT GTAGGGAATC GTGAGTTGCT
41251 TGGCTTATAC TGTGTTCTTG GTGTTAAGTA TTAACAGGAA TATGGCCTGG
41301 TAATTAGAAC TTGTCCATCA GAATTGCCAA AAGTGGGATT CGGGGGTCTC
41351 TGCCTATGGA GGATGTGGTT CAGAAATAAA GAATTTGAAT AGGATAAGCT
41401 GTAGGAGGAT CTTAGTATGA GAATGAGTAT CTGAAGATTA GCTGTGAGAG
41451 AGGGCAGAGC GATGGAGGGA ACAATGTGGG ACAGTGTGAA GCATGTGATC
41501 CAGGGGCCAT AACTTTTTTT GTTACTATTT TTTTAAATCA GAAACTTAGA
41551 TTTCACTGTC CTTTCTATCA AAGAAAAGGA CAAAAGATAA ACGTTCAAAA
41601 TTGGAATTTA TTTTCTTTT GGCAAATGTT AAATCTCACC TCTAATGAGA
41651 AATCATAGCT AATTAGGAGA TAACCTACAT GTAAGCATTT AGATTCAGTG
41701 CCATTAGAAG TGCTGGGTGG GTGATATCTG CAGGAGAAAA AAATGATGCT
41751 AGTTTTAAAA ATCTCTACTA TTACCGTGAA ATATTTTTAA ATGAAAACTT
41801 TCGTCTCTA AATATGACTG TGGAAAAGAA AATGAGTATA TTTAATAACA
41851 TCTTTTGACA TCTCTAGTAG TAACAGTAGG TCATCTTATT CATAAACCAA
41901 AATTTTACCA AATTTAGGCG CAGGCGCAGT GGCTCATGCC TGTAATCCCA
41951 GAACTTTGGG AGGCCGAGGC GGGCGGATCA CCTGAGGTCA GGAGTTAGAG
42001 ACTAGCCTCG CCAACATGGC AAAATCCCAT CTCTAGTAAA AATACAAAAA
42051 TTAGCCAGGC GTGGGGGCCC GTGCCGTGAA TCCTAGCCAC TTGGGAGGCT
42101 GAGACAGGAG AATCGCTTGA ACCCAGCGGG CAGAGGTTGC AGTGAGCCGA
42151 GATCGCGCCA TTGCACTCCA GCCTGGATGA CAGAACAAGA CTTTGTCTCA
42201 AAAAAAAAAA AAAAAAAAAA AAAAAAATTA ATCAAATTTT AAAACCAGGT
42251 TTTGTAGTAC ATTTAAATTG CATATTCCAA AGCAGTTGGG TTTGCCTGCG
42301 TTGCAGTTTA ATATTAAGCT ATACTTCCCT TTCAAATAAG GTATTTTCAT
42351 CGTTAAGCCT GTAAATTCTA GTTTGTCATT GTTTAGATAT TTATAGTCAT
42401 TTTAATATAT CTGTTTACGG CCAGCTGCAA TGGCTAACAC CTGTAAACTC
42451 AGCACTTTTT GAGGCCAAGG TGGGCCGATT GAGCTCAGGA GTTCGAGACC
42501 AGCCTGGGCA ACATAGTGAA ACTCCATCTA TACAAAAAAT CCAAAAAAAA
42551 AAAGACAGGT GTGGTGGCAT GTGCCTGTAG TCCAGCTAT CCCGGAGGCG
42601 GAGGCGGGAG GATGGCTTGA GCTTGGGAGG TCGAGGGTGC AGTGAGCTGT
42651 GATTGTGCCA CTGCACTCCG GCCTAGGTGA CAGAGCAAGA CCCTGTCTCA
42701 AAAAAAAAAA TCTCTTCACT CCTTAGCAGT GGTATTTTGT TAGCTAGAGT
42751 TGTCTACTA CTCTTTTGTG ATTTGTCTGT TAGGTCAAGG ACATGTTTC
42801 TGTTTATTCC AGAACTATAT TATCGAACTA TATTATCAGT CTTTCAAATG
42851 TCTTTTTTAG AGTCCTTCAA TAATGTTAAA CAGTGGCTGC AGGAAATAGA
42901 TCGTTATGCC AGTGAAAATG TCAACAAATT GTTGGTAGGG AACAAATGTG
42951 ATCTGACCAC AAAGAAAGTA GTAGACTACA CAACAGCGAA GGTATGTTTA
43001 AAGTTTAATT TTCATCTGTA ATTTGAAGGT GTTGAATTAT STATGGGTTT
43051 TGCAGTAAAC ATAAGGCCAC AGCCTTTTAA AAATATGTGC ACTAGAATAC
43101 TGTGACAGTG ACAATTTGTG TAGCATCTGT TTGGATCCAA TGAAGTTAGT
43151 TCCTCACGCT CCATTATGGA TGGTAGAAAT GCAGTAAGAA TTAGTGAAAA
43201 AGATTTTTCA GTGTTAATTG TGCCTCATT TCTCTTAGG AATTTGCTGA
43251 TTCCCTTGGA ATTCCTGTTT TGGAAACCAG TGCTAAGAAT GCAACGAATG
43301 TAGAACAGTC TTTTCATGAC ATGGCAGCTG AGATTAATAA GCGAATGGGT
43351 CCCGGAGCAA CAGCTGGTGG TGCTGAGAAG TCCAATGTTA AAATTCAGAG
43401 CACTCCAGTC AAGCAGTCAG GTGGAGGTTG CTGCTAAAAT TTGCCTCCAT
43451 CCTTTTCTCA CAGCAATGAA TTTGCAATCT GAACCCAAGT GAAAAACAA
43501 AATTGCCTGA ATTGTACTGT ATGTAGCTGC ACTACAACAG ATTCTTACCG
43551 TCTCCACAAA GGTCAAGAGT TGTAAATGGT CAATACTGAC TTTTTTTTTA
43601 TTCCCTTGAC TCAAGACAGC TAACCTCATT TTCAGAACTG TTTTAAACCT
43651 TTGTGTGCTG TTTTAAAAA TAATGTGTGT AATCCTTGTG GCTTTCCTGA
43701 TACCAGACTG TTTCCCGTGG TTGGTTAGAA TATATTTTGT TTTGATGTTT
43751 ATATTGGCAT GTTTAGATGT CAGGTTTAGT CTTCTGAAGA TGAAGTTCAG
43801 CCATTTTGTA TCAAACAGCA CAAGCAGTGT CTGTCACTTT CCATGCATAA
43851 AGTTTAGTGA GATGTTATAT GTAAGATCTG ATTTGCTAGT TCTTCCTTGT
43901 AGAGTTATAA ATGGAAGAT TACACTATCT GATTAATAGT TTCTTCATAC
43951 TCTGCATATA ACTTGTTGGT GCAGAATATT GTAATTTGTT GCACACTATG
44001 TAACAAAACA ATTTGAAGATA TGTTAATAA ATATTGTACT TATTGGAAGT
44051 AATATCAAAC TGTATGGTGA TAAGTATTGT TTTGATTCTT ATGGTTAAAG

FIGURE 3, page 14 of 21

```

44101 GGAAATAGAG CCTTGCATTA TATTCAACAC AGCCATTTGT GTGTGCACAA
44151 TGCAAACTAA GGTATTCTAG ACCTATCTTA GAGCAGCATC CAGTATTTCG
44201 TTTCTAGATA ATATGCCCCA TAACATGACC TAGAGGGGCT TCTGTGCTGT
44251 GTAGGGATT T AACCAACTTC AGTGGTTCAG GGAGCTCAA CTATATGTAA
44301 AACAAAGTTA GAATGTATGC TATCTAGCCC GTTATCTCTG ATCCTTCTCT
44351 AAAACCATT GAAATAGCTT CATTGATCAA CATTTCATAA ATGCATCTGT
44401 GGTAGAGGTA GAAAGCAGCA CCTTTCCTAA TTGGCAAATG ATCAGACTAA
44451 TGTGTGCTAA TGTTTTTCTT CCATGCTTTC AGTCAGATTC AACTATTTTA
44501 TCCTCCACAG TTGCTTAACT TGGTGTGGA GGAGGGTTTA AGCATTAAAG
44551 TAGGAAGCAG GAAATTTGAT TGCTCTAAAT TTAGAAATTA TATCCCTAAA
44601 AATTAACAACA TGAATACTGG GTGGTAATGA TAATTGAGGC AAATGTATTT
44651 ATTTTGGTGA CATTTTGCAT ATATGAAGAT TTTCTGAAAT AGGACCTTCA
44701 AGATCCTAGG GGGTTTTGTT TGGTTTTTAA TTGTGAGGAA TAAAAAATCT
44751 TCTGCCCCACA CTGGCATTTT AAGGTGACTG AGGTCAAACG TTGTTTCCTT
44801 AGGTTGAAAT AGCAGCCAAA ACATTCTTCA CGCAGGGGCT TGGGATATGG
44851 CTGCTGGCAA CACATTTTGT TGTGGGCTCC TTAATTTAAT GATAAAATTT
44901 AAGCTAAACA CAAGCCAAAA ATGAATAGGT TTTTAAAT TTTATTTTTC
44951 ACTAAACAGG CAATTGAAAT ACATGGTACA AAAATAAGTG GTAAGATAAT
45001 TGTAAAATGA AATGGACAGA ATATCAATT TTCCATCTAT GAAAATTTCA
45051 CAATAAAAAT CATAGTTTAC TTTGTATTAT AGGCGTGCTT GGTGGATCTA
45101 TTCATCCTCA CATAAGGCAA CTGACAAATT CCTGAAGTTA CCAATAGTTA
45151 TTTTGGTGAA GATCTTTAAT GCTTCAGAAG TTTTGT TTTT GCCTTAATAC
45201 AGTATAAAGG GGGAAAGAGT TCAGAACTA TTTTCTAAAG TAGCTAAATG
45251 ACACAAAACA AATGTCAAGA TACTGTGATG CCATGCCGTG CACTTCATTT
45301 TTACACAGTA AAAGTTGTTT AAATTGTCAG CTTATTCTTG GTGAGTTAGC
45351 GGAAACATTA CATGAACCTA AGATGAGCAT ATTTACAGAC TTAAGTTTGG
45401 AAAATTCAG CGTTCCTTTC CCCATGGCAG TAAAGATTGG GATTTACAAC
45451 AAATTCAGC ATGCCCTAAG ATTTGCTTCT ATGTATACGC CAATAAATGT
45501 GGTTCCTGAA AAAATATATA CCCCTTTATA CCCCATT TTT CAAGTACAAA
45551 CGGTTCAAAG CTACTACAGG TTTTAATAAT CTGTTCACT AGTAAAGGGA
45601 ATTACCACTT GTTCTAAATA TAAGGTGCTG CCATAAATTA GTTTACATAG
45651 TGAAGAAGAG TGTCTTAAA TCTAAGCAGC TGCACACTCT GTGAAATCCT
45701 TTCAGAAATGA TAGTCATTGT GGTCTGAGCA GTAATTTCTT ATTCTTCGAC
45751 CTTGGATTGA ATTTCCCTTA GCCTACATCT TGCCTTTCCA GCATATCTTA
45801 CCTCAAACCT TCTTTGTGTT CCATTCCAC CTAAGCTTCA AAATAGCCCT
45851 GTGTTGACGT CGCTTCCAT TTGCTGAGCT TACCTATGGA TCTCCAAGAA
45901 CCCAGATCTT GAACTGCTG ATCCAGCTTT GAGTATCATC ACTTCCCTGT
45951 GGATTTAACT TCCATTAATT TTAAGGGACT ACTAAGTTAT TCCAGTGTGG
46001 CATCACAGTG CAGTTAGCAA GCTCAGCTAC TTGACTCTAA TTTGGCCATG (SEQ ID NO:3)

```

FEATURES:

```

Start:      2181
Exon:       2181-2203
Intron:     2204-27090
Exon:       27091-27163
Intron:     27164-33853
Exon:       33854-33949
Intron:     33950-42859
Exon:       42860-42991
Intron:     42992-43239
Exon:       43240-43434
Stop:       43435

```

CHROMOSOME MAP POSITION:

Chromosome 2

ALLELIC VARIANTS (SNPs):

DNA				Protein		
Position	Major	Minor	Domain	Position	Major	Minor
397	T	-	Beyond ORF(5')			
2326	A	G	Intron			

3486	C	A	Intron
6651	-	A	Intron
8190	T	-	Intron
8281	T	C	Intron
11546	A	G	Intron
11670	C	T	Intron
11688	A	G	Intron
14938	A	C	Intron
22261	G	A	Intron
22852	G	A	Intron
27253	A	C	Intron
28098	-	A	Intron
28597	G	T	Intron
31431	C	T G	Intron
35704	C	T	Intron
35728	C	T	Intron
36690	C	T	Intron
41002	G	C	Intron
41033	A	G	Intron
43161	C	T	Intron
43765	A	G	Beyond ORF(3')
44713	G	T	Beyond ORF(3')
44831	C	T	Beyond ORF(3')

Context:

DNA

Position

397 TGCTCTGTGCGCCAGGCTGGAGTGCAGTGGCCTCTCGGCCACTGTAGCCTCCGCTCCC
GGGTTCAAGCAATTTTCTGCCTCAGCCTCCCGAGTAGCTGGGATTACAGGCACGCGCCA
CCATGCCTGGCTAATTTTGTATTTTGTAGTAGAGACAGTGTTCACCATGTTGGCCAGGC
TGGTCTTGAATTCCTGACCTCGTGATCTGTCCGTTTTGGCCTCTCAAATTCCTGAGATTA
CAGGCATGAGCCACCGAGCCTGGCCAGTTTTCTGAGTTTTTATTTGAAATCAAAATAAGC
[T,-]
TTTTTTTTTTTTTAAATGGGCTTTAGAGTCCAGGGTAACGAACACTTTTTGGTGCCTATT
ACTGAACCATTCAGGGTATTCTGGGGTGGTGACCGTGTTCATTTCAGAAACCAACATGT
TCATTTTCAAGAACCAAACTCGGGTAACTTTGTATAAGTTCATCAACTAAGGCCCATGGCA
GAATTTGAGGGCTAAGGGGTGAATTAGTGTATGGGTAGAAATAAGTGCCTTCTTCTAT
ATTTTGGCGTTGTAGGAATTTAAAGTGATTCTGCAGTAAGTCTCAGGAGACAATTTCTT

2326 GCTGATTGTGTTCTAGGGGACGGAGTAGGGGAAGACGTTTGCTCTCCCGGAACAGCCTAT
CTCATTCTTTCTTTTCGATTACCCGTGGCGCGGAGAGTCAGGGCGGCGGCTGCGGCAGCA
AGGGCGGCGGTCGGCGCGGCGGAGCTGCAGTGACATGTCCAGCATGAATCCCGAATAGT
GAGTTCAGGAGAGCACCGGTGCGGCTGGGTCCGTGGGCCAGCTTGGGGGATCTTAAAGGGG
TCGAGGAGGGTTGGGGCAGAACTCGGGGCATCGGCTGGGGTGGGCGAGGGTGATGGGTC
[A,G]
GGAGAGGCTGGCGGCGGGAGTGGGCCCCATTGTCTGACGCGGAGGGGCGGCGCGCGG
GGGAGGGGTGGGCGCGGAGGGGTGAGCCGCCCGGGCCTGGACCGGGTCAGGTTAGAGGGC
CTGACTGCGGGGCGGGTGCTGAGGAAGCCTGCCGAGGGGCGTGGGGCGGTGTGAAGGGGT
ATCTTCTCTCGGAGGCAGTGACTTTTGAAGGAGGACTTGTCTCTAAGGGGAGGGGATGGG
GTGGGAGAGCCCTTCTAGAGGCACTGTGACACCCTGCGCCGCACTCTGCGGAGCTGTC

3486 CTGGGAAGTGGTGTCACTTCCCTTGGGTAGAGTTTGTGGGCTCTCCTCAATGGCCCTT
TAAAAATTTCTCTACAGTTTACATGCATGTAAAGTAATGAATAATTGGAAGAGACCGAA
TTGGTATTCTTTTTCAGTGTCAAAGGCCCTTGGAGGATGGGGGAAAATCAGTATTTGTTG
TAAAGTTGAGTTTATTTGCTGGTTTGGTCAATTACTGCTAGACATTTCCCTTAAAGG
TCCACCCACCAGTTTAGCTGACTGTGATGTGTGTGACATGGCTCTTGCAAATGCTTA
[C,A]
AAGTTTGTAAATAGTGTGGCTTGAAGCTGAAATCTTTTGCACTAAACAGAAACCGTAGTA
TTTTATTAGAATTTTCATGCTTTAGAAGTTGAGGGTAGTGTCTTGTAGTGACATTTGCTG
TGTTGACAGTTTAAAAAATTTTTTTTCAAGGGCTCCAAGGACAAAGTTGGTTTTGCAC
AGTTGAACGGAGTGAACCTTGAGGTTCTTAATTTAGTAGTTTTCTTGTTAACAATAAAGA
ACATGGATTACTGCTTTATCGAGGTTTATAGACCTCTACTGTTTCAGGAAATTTCTGAA

FIGURE 3, page 16 of 21

6651 TTTCAGCACATTAAGAAATGCTTAACATGGCCAGGCGCAGTGGCTCACGCCTGTAATTCT
CAGCACTTTGGGAGGCCGAGGTGGGCGGATCATTGAGGTGATGACCAGCCTGGCCAACA
TGATGAGACACTGCCTCTACTAAAAATACAAAAATTAGCTGGGTGTGGTGGTGCACGCCT
GTAATTCCAGCTACTCAGGAACCTGAGGCAGGAGAGTCACTTGAACCTGGGAGGCGGAGG
CTGCAGTGAGTCCAGATCATGCCACTGCCTCCAGCCTGAGGGACAGAGTGAGACTCCTC
[-, A]
AAAAAAAAAAAAAAAAAAGAAAGAAATACTTAACATTATTCTCGTGATTATTCTCATAAC
ATTTTTCATAATCCACTGGCTTCCAGTGGATTTTTTTAGTGTCAAGAAAATAATTTTGAT
TGGTTCATCTTTAAGGAATGTGTTAAGAATAAAGCATGTCTACCTGTCTTCAGTATACCA
GCTAACTATAGTAGGAAGAAATATAGTAGTCTACTTAGATCAACTATAATTCTTTAATGC
AGAAAAAGTTTAAAGTATTACCTTATTTTGTAGCCCCATCCCCTTAAGTATATCATGGC

8190 AGACCGGCCTGGCCAATGTGGTGAAACCCTGCCTCTACTAAAAACACCAAATTAGCTAGG
CGTGGTGGTGTGCGCTTGTAGTCCCAAGCTACTGAGGAGGCTGAGACAAGAGAATCGCTT
GAATCTGGGAAAAAGAGGTTGCCGTGAGCCAAGATTGGCCACTGCCTCCAGCCTGGGTG
ACAGAGTGAGATTCTGTCTCAAAAAATAAAAAATAAAAAATTTCCCTTTAATCAAATT
AAGTTAAATGAGGGATGTTAGACAGTTTTTAACCATCAAATATTTAGTTTAGTTTTT
[T, -]
TTTTTAACGTTGTCTTAAAGATGGAAGTGCTTCAAAATCAAATCTTCCTTGCCAGTTCTC
TACTTGGCTTCTTTTTTTCTTTTTTGAGATAGAGTCTCACTTTGTCACTGGAGTGCGTT
GGCGTCTCACTCGGCTCACTGCAACCTCCGCCTTCCAGGTTAAGTGATTCTTCCACCTCA
GCCTCTCAAGTAGCTGGGAGTACAGGTGTGTGCCACCACACCCGGCTAATTTTTGTAGTT
TTAGTAGAGACAGGGTTTCACTATGTTGGCCAGGCTGGCCTCAAACCTGACCTCGTGA

8281 CTGAGGAGGCTGAGACAAGAGAATCGCTTGAATCTGGGAAAAAGAGGTTGCCGTGAGCCA
AGATTGGCCACTGCCTCCAGCCTGGGTGACAGAGTGAGATTCTGTCTCAAAAAATAAA
AAATAAAATTTCCCTTTAATCAAATTAAGTTAAATGAGGGATGTTAGACAGTTTTT
AACCATCAAATATTTAGTTAGTTTTTTTTTTTTTAAAGTTGTCTTAAAGATGGAAGTGC
TTCAAATCAAATCTTCCTTGCCAGTTCTCACTTGGCTTCTTTTTTTTTCTTTTGTAGA
[T, C]
AGAGTCTCACTTTGTCACTGGAGTGCGTTGGCGTGATCTCGGCTCACTGCAACCTCCGCC
TTCCAGGTTTAAAGTGATTCTTCCACCTCAGCCTCTCAAGTAGCTGGGAGTACAGGTGTGT
GCCACCACACCCGGCTAATTTTTGTAGTTTTTAGTAGAGACAGGGTTTCACTATGTTGGCC
AGGCTGGCCTCAAACCTCCTGACCTCGTGATCCACCACCTCAGCCAAATTGCTGGGATTA
CTTGTGTGAGCCACGCGCTGGCTTCACTTGGCTTTTAAAGGAATTTGCTTTCTGAG

11546 GTTACATTTAACCATTATGCGTGTAGCCATACTCACGTTACATTTGATGCATCTGC
TCCTTTGTGTCTATATACTCATATAACATTTTGCATAAAGTTATAGGCAGTTACACCA
AGGCTGTTTCATGAACCTCAGATTAAAGATACTTGATTAGGAGATTGAAAACAGAAAAGA
GAATGTTAACTATCATTATCAATATTAATGTGAAAATCTGAGAGTGACAAAGCTTAGC
TTTAAATCTGGTATCCCAAACCTCATTGAGTTTTTTTTTTTTTTTTTTTTTTTGTAGAC
[A, G]
AGGTGTGCGCTTGTCCCCAGGCTGGAGTGAGTGGTGTGATCTTGGCTCACTGCAACCT
CCACCTCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCTGAAGTTGCTGGGATTACAGG
CTGCGCCACCACGCCCAGCTAATTTTTGTATTTATAGTAAAGACGGAGTTTACCTTAT
TGGCCAGGCTGGTCTCAAACCTCCTGATCTTGTGATCCTCCCGCCTCGGCCTCCCAAAGTG
CTGGGATTACAGGTGTGAGCCACTGTTCCCGGCCTAATTTGAGTTTAAATGTGGAGTT

11670 TGTTTCATGAACCTCAGATTAAGAATACTTGATTTAGGAGATTGAAAACAGAAAAGAGAAT
GTTAACTATCATTATCAATATTAATGTGAAAATCTGAGAGTGACAAAGCTTAGCTTTA
AATCTGGTATCCCAAACCTCATTGAGTTTTTTTTTTTTTTTTTTTTTTTGTAGACAAGG
TGTCGCTTTGTCCCCAGGCTGGAGTGAGTGGTGTGATCTTGGCTCACTGCAACCTCCA
CCTCCAGGTTCAAGTGATTCTCCTGCCTCAGCCTCTGAAGTTGCTGGGATTACAGGCTG
[C, T]
GCCACCACGCCCAGCTAATTTTTGTATTTATAGTAAAGACGGAGTTTACCTTATTGGC
CAGGCTGGTCTCAAACCTCCTGATCTTGTGATCCTCCCGCCTCGGCCTCCCAAAGTGCTGG
GATTACAGGTGTGAGCCACTGTTCCCGGCCTAATTTGAGTTTTTAAATGTGGAGTTAAG
ATGTTAGTCTTAAAGTGGGTAGATGAAATTTATAAAAATAGTCAAATAGCTAAATTTAT
AAAAGGCCATTTGAAACAATTTTGTGAAATATATAATGTGGATAATTATGTAGTGCTTTA

11688 TAAGAATACTTGATTTAGGAGATTGAAAACAGAAAAGAGAATGTTAACTATCATTATCAA
TATTAATGTGAAAATCTGAGAGTGACAAAGCTTAGCTTTAAATCTGGTATCCCAAACCT

FIGURE 3, page 17 of 21

CATTTGAGTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGACAAGGTGTCGCTTTGTCCCCAG
 GCTGGAGTGTAGTGGTGTGATCTTGGCTCACTGCAACCTCCACCTCCCAGGTTCAAGTGA
 TTCTCCTGCCTCAGCCTCTGAAGTTGCTGGGATTACAGGCTGCGCCACCACGCCCAGCTA
 [A, G]
 TTTTTGTATTTATAGTAAAGACGGAGTTTCACCTTATTGGCCAGGCTGGTCTCAAACCTC
 CTGATCTTGTGATCCTCCCGCCTCGGCCTCCCAAAGTGTGGGATTACAGGTGTGAGCCA
 CTGTTCCCGGCCTAATTTGAGTTTTAAATGTGGAGTTAAGATGTTAGTCTTAAAGTGG
 GTTAGATGAAATTTATAAAAATAGTCAAATAGCTAAATTTATAAAAGGCCATTTGAAACA
 ATTTTGTGAAATATATAATGTGGATAATTATGTAGTGCTTTATGTGTAGATTGGTGGTTA

14938 CATGGTAGTGTGCACCTGTAGTCCCAACCACTTGGGAGGCTGAGGTGGGAGGATTGCCTG
 AGGCCAGGAGTTTGAGACCTGGGCAGCATATGAAGACCTGTCTCTAAAAAACTAAAAAT
 AAAAAATAGCCAGGTGTGGTTGGTGTGCTTGTGGTCCCAGCTACTCAAGAGGCTGAGGCA
 AGAGGGTTGCTTGAGCCCCAGAAGTTGGAGGCTGCCGTGAAGTGTGATTGCACCACTGCAC
 TTCAGCCTGGGTGACATAGCAAGACCTGTCTCTGTGGTGGTGGTGGTGGGGGTGGGGG
 [A, C]
 AGGGATTTAAGAAGGGTTTGTGAGGTATGTATTATTTATAAATGGGCTTTTAACCTTACC
 CTTACACATCTGGGTTGAAATTAATTGTATCCATTCTCAGTTTTTCTGTCTTGCTATATA
 TTTAACTTGGAGACTTAGAGGTATGGATGTCTTCTATGAAAAGCAAATGAAGCAGAG
 GGCTGCCTTCTCTGTCTGTAGAGGGCACACTTGCTGCAGAGCATGTTACTGTTTTATGCA
 TTGCTAGGCTTTGGGAGTTGTGACTTGTATGATCATAGTACTTACAACATTAGTTGGCA

22261 CACCCACAGATAGCTATGTCAAACGTAAGGGTGGAGAAACACAGACCCCAAACCTTCTCGA
 GGGTAGAAAATATGAGGTTATAGTAGATTAGAAGTACAAAAAGCTAGAGGAAGTTCTGAA
 CTGGAAACAGTGGATAGGATTTACTAGAATAATTTACGAGGGTGACAATTGTAAATCTTC
 ATAGGTTTCTTTTTTTTCTTCTCTTTTTTTTTTTTGGAGATGGAGTCTCGCTCTGTTG
 CCCAGGCTGGAGTGCAATGGCGCAGTCTCTCTCACTGCAACCTCCGCCTCCTGGGTCCA
 [G, A]
 GTGATTTCTCCTGCCTTAGCCACCCAAGTAGCTGGGATTACAGGCATCTGCCACCATGCTG
 AGCTAATTTTTGTATTTTTTTTTTTTAGTAGAGACGGGGTTTACCATTGTTGGTCAGGCTG
 GTCTTGAAGTCTGACCTCAGGTAATCCACCCACCTTGGCCTCCCAAAGTGTGGGATTA
 CAGGTGTGAGCCACCGCGCCAGCCAAATTTTATTGGTTTCTAACTAGCGTAATTTAG
 TTTTTTTCACCTAAGTCAAATATATATTATGTAGGATAAAAACCTTAGTGATCCAAATTC

22852 ATCCAAATTCATGAGGAATGAAGAATAAATACATTTAAAGTCTTACCATTGCTAAATTA
 GTCTTGGCTCTTTGTACCAAATTTCTGTCTTGTGCTCTGTAATTTATATTTGTATATT
 TTCTATCAACATTTTTACTGTGTGGTGTTTGTAAATTATAAAAAAGCTTTTAAAGCAAAC
 TCAGAACAATGAATTTCTCAGGAATATTCAGTATATTTACAGTTGAGAAATAAACTACTTC
 TGTAGTAGGTAATTTAAATGTCCCAATGCAAGTTAACGTGTCACTGATCACGCTATTCA
 [G, A]
 GTGTGTGTCTTTGATAAGGGGAGGTGGGGAAGTTTGTGGGTTTGATTTTATTTGCCTTTC
 TCATGTGACTGTTGTATGTTAGTAAACAAATGGTTTGCAGAGAACCAAGTAGTCTTTTG
 CAAAGATTGTCTTATACAGAGCACTCAATTTCTTATATTATTTATAATGGCTTTAATTTA
 AGCCTTAAATTATTAGAACTCATAAATAATTTTTTTTATTTGTTTGTGATGGAGTT
 TCGCCCTTATTGTCCAGGCTGAAGTACAATGATGTGATCTTGACTCACTGCAACCTCCGC

27253 GCTTAAGCCATGCATGGGCTTTATAGGAGATGTAGTCTTACAGTGAGTTGTTATTTGTA
 GCTGTGTTTTTGTTTTTGTATAGCTTATAGCAATGCAGTGTGCTTTTTATTAACATCATT
 TTCTTTTTCTTTTTTGCAGTGATTATTTATCAAGTTACTTCTGATGGCGACTCAGGGGT
 TGGAAAGTCTTGCTTCTTCTTAGGTTTGCAGTAAGTTGAAATTGAAATGTCTTTACAAT
 TAATGGTACAATTAATGCTATGTATGTTTTCTAGGTAGATAAAATTAACAGTTTATTTC
 [A, C]
 GAATAAGTTAATTTCTCCAGAATTTATATATTTAAAGACTCCAAATATACATCCCCAGTG
 GTATCTTGGACTGTTAAATAGAAAAATATTGTTGCTCTTAAAGAAATTCAGTGAAGTCT
 GGTATAAAGTCAGAATGTCTAATACTTTTGGTCAGAGTCAAACAGCAGTTCCAATATAG
 GCAGCAAGTTAAAGGGGTAGTTGGTGGCCTGTGTTGAAAGCGACTTGATGAAATAAATC
 TTTAAATTAACCTTTAGTAGAATAAAAAGAAAAAGCAGAGCCAGGTGACGCAGTGGAATCA

28098 CTTTAAATTTAGCATGTTTCTGGCCAGGTGCGGTGGCTCACGCCTGTAATCCCAGCACT
 TTGGGAGGCCGAGACGGGCGGATCACAAGGTCAAGAGATTGAGACCATCCTGGCTAACAC
 GGTGAACCCCGTCTCTACTAAAAATACAAAAATCAGCTGGGTGTGGTGGCCACACGCCCT
 GTAGTCCCAGCTACTCGGAGGCTGAGGCAGGAGAATCGCTTGAACCCAGGAGGCGGAGG
 TTGCAAGTGAGCTGAGATGGTGCCACTGCACTCCAGCCTGGCAACAGAGCAAGACTGTCTC

FIGURE 3, page 18 of 21

[-, A]
 AAAAAAAAAAGAAAAAAAAAATAAAAAACAAATTAGCATGTTTCCCTTCTAGAGATCATTGT
 TTCTCAGAGCATGGACCAAAGACTCCTGGGGTTACCAAGACCCTCTCAGGTAGCCCATG
 AGGTCAAATATCCTAATAATACTAAGATGTTAGTATTTGTAAGGAAATATTTACTTGGT
 AATAATACTAATAATAAAGATGTTTGCCTTTTTCAGTGATGACATTGGCTCTGGTACAAA
 AGCATGTGGGTAAAATTGCTGCTGGCTTGGTACACATCAAGGCAGCGCTAAGCTCCAAAT

28597 GATGTTTGCCTTTTTTTCAGTGATGACATTGGCTCTGGTACAAAAGCATGTGGGTAAAATTG
 CTGCTGGCTTGGTACACATCAAGGCAGCGCTAAGCTCCAAATTGTACTCATGGTGATGGC
 ATTCTTTACCTCTGTGCCCTCACAGGAACAAAAACAAGCCGTGCCATTTTTATTGAAGAT
 TGTCCTTGACAAAACAGTTAAAATGATTAATTTTTGAAAAATGTTGATCCATGAGTATTC
 CTTTAAAAATATTTGTGAAGAAATGGGAAGTTCACATAAAACAATGTTTTTTTTTTGTTT
 [G, T]

TTTTTTTTTTTTTTTTTGGAGACAGATTCTGGCTGTGTTGCCAAGGCTAGAGTGCACTGGC
 GTCTGGCTCCAGGCTCAAGCTGTTCTCCCACTTCAGCCTCCCAAGTGGCTGGGACCTCC
 CAAGTGGATGCGCCATCATGCTTGGCTGATTTTTGTATTTTTTGTAGTGACAAGGTCTC
 ACTGTGTTGCACAGGCTGGTCTCAAACCTCTGAGCTCAAGCGATGCATGTGCCTCAGCCT
 CCCAAAGTGCTGGAGAAAGCACTTTTTACTGCATACTGGCTAGTGTGTTGTTATTTTGG

31431 CTGCATTTTTTTTTTTTTTTTTTGGTTTGGAGATGGAGTCTCGCTCTGTGCGCCAGGCTGGA
 GTGCAGTCGTGCAATCTCGGCTCACTGCAGCCTCCACCTCATGGGTTCAAGCGATTCTCC
 ATCTTGGTCTCCTGACTAGCTAGGTTTACAGGCGTGTGCCATCACACCCACTAATTTTTT
 GTATTTTGTAGTAGACAGGTTTACCATGTTGGCCAGGCTGGTCTTGAACCTCCTGATC
 TAAAGTGAGCCTCCACCTTGGCCTCCCAAAGTGCTGGGATTACATATGTGAGCCACTGC
 [C, T, G]

CCTGGCCTCTATATACTTCTATAGTACCTGATACTTATTAGGCACTCAATTACAACATAA
 CTTTTTTTTTTTTTTTTTTTTTGGAGACAGAGACATGCCTTGTGCGCTGGGCTGGAGTGC
 AGTGGCACAGTCTCGGCTCACTGCAACCTTCACCTCCCGGGTCAAGTGATTCTCCTTCC
 TCAGCCTCCCGGTAGCTGGGATTACAGGCGCCCGCCACCACGTCCAGCTAATTTTTGT
 ATTTTTAATAGAGATGAGGTTTACCATCTTGGCCAGGCTGATCTCAAACCTCCTGACCTT

35704 ATGTGTGATCATTTGGTGTTTATAAGATTTGGGTGTGTATTCTGTGTGAAACATTCATAT
 TTTGTTACTTTCCTGTGGCTGGAAGGGATCTTATAGGACACTGTCTTTCATCTTGTCTG
 TCTTTCATCTTAAATAGGAATTTCTTTTCCATGCCTGAAGGCCTCATTTTGAACATTTG
 TTTGTTTGTTTTTTATTTTTTGGAGATACAGTATTGCTCTGTCTCCAGGCTGGAGTGCA
 GTGGCGCGATTGAGCTCACTGCAACCTCCGCCTCCTGGGTTCAAGTGATTCTCCTGCCT
 [C, T]

AGCCTCCCTAATAGCTGGGATTACATGTGTGTACCACCATGCCCGGACAATTTTTTTTTT
 TTTGAGATGGAGCCTTGCTTTGTGCGCCAGGCTGGAGTGCCAGTGGTGCAATCTTGGCTC
 GCTGCAGCCTCCGCCTCCCAAGGTTCAAGCAGTTCTCTTGCTCAGCCTCCTGAGTAGCTG
 GGATTACAGGCGTGCGCCACCACACCTGCTAATTTTTTGTATTTTTAGTAGAGACAGAG
 TTTACCATGTTGTTAGGCTGGTCTCGAACTCCTGACCTCGTGATCTGCCTGACTCGGC

35728 GATTTGGGTGTGTATTCTGTGTGAAACATTCATATTTTGTACTTTCTGTGGCTGGAA
 GGGATCTTATAGGACACTGTCTTTCATCTTGTCTGTCTTTCATCTTAAATAGGAATTC
 TTTTCCATGCCTGAAGGCCTCATTTTGAACATTTTGTGTTTGTGTTTTTATTTTTTGA
 GATACAGTATTGCTCTGTCTCCAGGCTGGAGTGCACTGGCGCGATTGAGCTCACTGCA
 ACCTCCGCCTCCTGGGTTCAAGTGATTCTCCTGCCTCAGCCTCCCTAATAGCTGGGATTA
 [C, T]

ATGTGTGTACCACCATGCCCGGACAATTTTTTTTTTTTTTGGAGATGGAGCCTTGCTTTGTC
 GCCAGGCTGGAGTGCCAGTGGTGCAATCTTGGCTCGCTGCAGCCTCCGCCTCCAGGTT
 CAAGCAGTTCTCTTGCCTCAGCCTCCTGAGTAGCTGGGATTACAGGCGTGCGCCACCACA
 CCCTGCTAATTTTTTGTATTTTTAGTAGAGACAGAGTTTACCATGTTGGTTAGGCTGGT
 CTCGAACTCCTGACCTCGTGATCTGCCTGACTCGGCTTCCCAAAGTGCTGGGATTACAGG

36690 AAAAAAAAAAAAAAAAAAGTAACCAGGTGTGGTGGTCCATGCCTGTAGTCTAGCTCCCCAG
 GAGACTGAGGTGGGAGGAATGTTTGGAGCCAGGACTTCAAGGCTGCAGTGAGGCAAGATT
 GCACCATTCACCCAGCTTTGGGGACAGAGTGAGAGACCCTGTCTCAAAAACAAAATAA
 GGCTGGGCGCAGTGGCTGTCCGGGCGTCTGTTTACGCTTATAGTCTTAGCACTTTGGG
 AGGCCAAGGTGGGCAGATTGCCTGAGCTCAGGAGGTCTAAGACCAGCCTGAGCAACATGG
 [C, T]

GAAACCTCATCTTTGCAAAACATACAGAAAAAACAAAAAACCAAAACCTCTAGTT
 GCCAGTTATTTTTTTTTTATTATTCCTAGTGATTCTTCTTTTTTCTTTTTCTGAGACAA

FIGURE 3, page 19 of 21

AAATTTCACTTTGTCTCCCTCGCTAGAGTGCAGCGGTGAGCTCACTACATGATTCTTTTA
GAGACATGTTAATTCCTTATATTGAGCTGAAGCCTGTTTCTTTTACTTCTGTCTCTCTT
ATTCCTCCGCTTGTAGAGCTGCCTGAATCAGATTAATTCCTCTTTTATTGGCAAGCCTG

41002 GAGTTGAGGACTAATGTTTCTATATCACATCCTGATAATCTCCACAGTTATGAAAACATA
ACTATTTCCCCTCCCCTCCTACACTTTTCCCCAACTTTATTTTAATGGAATTGTTTGGATT
TCTTGATTGTTTTGTAAAGTGGGACACAGCAGGCCAGGAAAGATTTGGAACAATCACCT
CCAGTTATTACAGAGGAGCCCATGGCATCATAGTTGTGTATGATGTGACAGATCAGGTAA
GTTCCAAGAGGAGATTGTGTTACAGTGACCAAGTAGGAAGCCATTATTTGATTAAATGTCA
[G, C]
ATTCATTTACTACTTCATATATAAGCCATCAGTATTAATTTTATGGCAGAAAACTTTGTG
CACTCTCAAATATAAATGTGAATCACTTAAAAGACATTTGTTTTCTGTAATAAATAAAA
GATTAGTAATTAGTTTTACGTTTGTCTTCAAGGGATTCTGGTTGTATTTATTGTCAACTA
AATAACTTTGATCAAAATAGCCAAGACTCTAACATATAGGCAAGAGTTTGTAGGGAATCGT
GAGTTGCTTGGCTTATACTGTGTTCTTGGTGTAAAGTATTAACAGGAATATGGCCTGGTA

41033 CTGATAATCTCCACAGTTATGAAAACATAAATTTTCCCCTCCCTCCTACACTTTTCCC
AACTTTATTTTAATGGAATTGTTTGGATTTCTTGATTGTTTGTAAAGTGGGACACAGC
AGGCCAGGAAAGATTTGGAACAATCACCTCCAGTTATTACAGAGGAGCCCATGGCATCAT
AGTTGTGTATGATGTGACAGATCAGGTAAGTTCCAAGAGGAGATTGTGTACAGTGACCA
AGTAGGAAGCCATTATTTGATTAAATGTGAGATTCACTTACTACTTCATATATAAGCCATC
[A, G]
GTATTAATTTTATGGCAGAAAACTTTGTCCACTCTCAAATATAAATGTGAATCACTTAAA
AGACATTTGTTTTCTGTAATAAATAAAGATTAGTAATTAGTTTTACGTTTGTCTTCAA
GGGATTCTGGTTGTATTTATTGTCAACTAAATAAATTTGATCAAATAGCCAAGACTCTAA
CATATAGGCAAGAGTTTGTAGGGAATCGTGAGTTGCTTGGCTTATACTGTGTTCTTGGTG
TTAAGTATTAACAGGAATATGGCCTGGTAATTAGAACTTGTCCATCAGAATTGCCAAAG

43161 AGTCCTTCAATAATGTTAAACAGTGCGCTGCAGGAAATAGATCGTTATGCCAGTGAAAATG
TCAACAAATTTGTTGGTAGGGAACAAATGTGATCTGACCACAAAGAAAGTAGTAGACTACA
CAACAGCGAAGGTATGTTTAAAGTTAATTTTCATACTGAATTTGAAGGTGTTGAATTAT
GTATGGGTTCTGCAGTAACAGTAAGGCCACAGCCTTTTAAAAATATGTGCACTAGAATAC
TGTGACAGTGACAAATTTGTGTAGCATCTGTTTGGATCCAATGAAGTTAGTTCTCAGCT
[C, T]
CATTATGGATGGTAGAAATGCAGTAAGAATTAGTGAAAAAGATTTTTCAGTGTTAATTGT
GCCTCATTATCTCTTAGGAATTTGCTGATTCCCTTGGAATTCGTTTTTGGAAACAGT
GCTAAGAATGCAACGAATGTAGAACAGTCTTTCATGACGATGGCAGCTGAGATTAAAAAG
CGAATGGGTCCCGGAGCAACAGCTGGTGGTGCTGAGAAGTCCAATGTTAAATTCAGAGC
ACTCCAGTCAAGCAGTCAGGTGGAGGTTGCTGCTAAATTTGCCTCCATCCTTTTCTCAC

43765 AATGAATTTGCAATCTGAACCCAAGTGAAAAACAAAATTGCCTGAATTGTACTGTATGT
AGCTGCACTACAACAGATTCTTACCGTCTCCACAAAGGTCAGAGATTGTAATGGTCAAT
ACTGACTTTTTTTTTTATTCCCTTGACTCAAGACAGCTAACTTCATTTTCAAGACTGTTT
AAACCTTTGTGTGCTGGTTTTATAAATAATGTGTGTAATCCTTGTGCTTTCCTGATACC
AGACTGTTTCCCGTGGTTGGTTAGAATATATTTGTTTTGATGTTTATATTGGCATGTTT
[A, G]
GATGTCAGGTTTAGTCTTCTGAAGATGAAGTTGAGCCATTTTGTATCAAACAGCACAAGC
AGTGTCTGTCACTTTCCATGCATAAAGTTTAGTGAGATGTTATATGTAAGATCTGATTG
CTAGTTCTTCCCTGTAGAGTTATAAATGGAAAGATTACACTATCTGATTAAATAGTTTCTT
CATACTCTGCATATAAATTTGTGGCTGCAGAATATTGTAATTTGTTGCACACTATGTAACA
AAACAACGAAGATATGTTTAAATAAATATTGTACTTATTGGAAGTAATATCAAACGTAT

44713 AAGCAGCACCTTTCCCTAATTGGCAAATGATCAGACTAATGTGTGCTAATGTTTTCTTCC
ATGCTTTCAGTCAGATTCAACTATTTTATCCTCCACAGTTGCTTAACTTGGTGTGGAGG
AGGGTTTAAGCATTAAAGTAGGAAGCAGGAAATTTGATTGCTCTAAATTTAGAAATTATA
TCCCTAAAAATTAACATGAATACTGGGTGGTAATGATAATTGAGGCAAAATGATTTTAT
TTTGGTGACATTTTGCATATATGAAGATTTTCTGAAATAGGACCTTCAAGATCCTAGGGG
[G, T]
TTTTGTTTGGTTTTTAATTGTGAGGAATAAAAAATCTTCTGCCCACACTGGCATTTTAAG
GTGACTGAGGTCAAACGTTGTTTCTTAGGTTGAAATAGCAGCCAAAACATTCTTCACGC
AGGGGCTTGGGATATGGCTGGCAACACATTTTGTGTGGGCTCCTTAATTTAATGAT
AAAATTTAAGCTAAACACAAAGCCAAAAATGAATAGGTTTTTTTAAATTTTATTTTCACT
AAACAGGCAATTGAAATACATGGTACAAAAATAAGTGGTAAGATAATTGTAAAATGAAAT

FIGURE 3, page 20 of 21

44831

GGAGGGTTTAAGCATTAGATAGGAAGCAGGAAATTTGATTGCTCTAAATTTAGAAATTA
TATCCCTAAAAATTAAACATGAATACTGGGTGGTAATGATAATTGAGGCAAATGTATTT
ATTTTGGTGACATTTTGCATATATGAAGATTTTCTGAAATAGGACCTTCAAGATCCTAGG
GGGTTTGTGTTGGTTTTAATTGTGAGGAATAAAAAATCTTCTGCCCACACTGGCATT
AAGGTGACTGAGGTCAAACGTTGTTTCCTTAGGTTGAAATAGCAGCCAAAACATTCTTCA
[C,T]
GCAGGGGCTTGGGATATGGCTGCTGGCAACACATTTTGTGTTGGGCTCCTTAATTTAATG
ATAAAATTTAAGCTAAACACAAGCCAAAAATGAATAGGTTTTTTTAAATTTTATTTTCA
CTAAACAGGCAATTGAAATACATGGTACAAAAATAAGTGGTAAGATAATTGTAAATGAA
ATGGACAGAATATTCAATTTTCCATCTATGAAAATTTACAATAAAAAATCATAGTTTACT
TTGTATTATAGGCGTGCTTGGTGGATCTATTCATCCTCACATAAGGCAACTGACAAATTC